



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A61K 47/48, 39/00	A2	(11) International Publication Number: WO 98/36779 (43) International Publication Date: 27 August 1998 (27.08.98)
(21) International Application Number: PCT/US98/02945 (22) International Filing Date: 18 February 1998 (18.02.98) (30) Priority Data: 08/801,263 19 February 1997 (19.02.97) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/801,263 (CON) Filed on 19 February 1997 (19.02.97) (71) Applicant (for all designated States except US): UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL [US/US]; 308 Bynum Hall, Campus Box 4105, Chapel Hill, NC 27599-4105 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): JOHNSTON, Robert, E. [US/US]; 101 Marin Place, Chapel Hill, NC 27516 (US). DAVIS, Nancy, L. [US/US]; 132 New Castle Drive, Chapel Hill, NC 27514 (US). SIMPSON, Dennis, A. [US/US]; 19A Deer Mountain Road, Pittsboro, NC 27312 (US).		(74) Agents: MAGRI, Karen, A. et al.; Myers, Bigel, Sibley & Sajovec, L.L.P., P.O. Box 37428, Raleigh, NC 27627 (US). (81) Designated States: AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GE, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: SYSTEM FOR THE <i>IN VIVO</i> DELIVERY AND EXPRESSION OF HETEROLOGOUS GENES IN THE BONE MARROW		
(57) Abstract The present invention provides a method of delivering immunogenic or therapeutic proteins to bone marrow cells using alphavirus vectors. The alphavirus vectors disclosed herein target specifically to bone marrow tissue, and viral genomes persist in bone marrow for at least three months post-infection. No or very low levels of virus were detected in quadriceps, brain, and sera of treated animals. The sequence of a consensus Sindbis cDNA clone, pTR339, and infectious RNA transcripts, infectious virus particles, and pharmaceutical formulations derived therefrom are also disclosed. The sequence of the genomic RNA of the Girdwood S.A. virus, and cDNA clones, infectious RNA transcripts, infectious virus particles, and pharmaceutical formulations derived therefrom are also disclosed.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

-1-

SYSTEM FOR THE *IN VIVO* DELIVERY AND EXPRESSION OF HETEROLOGOUS GENES IN THE BONE MARROW

5

FEDERALLY SPONSORED RESEARCH

This invention was made with Government support under Grant Number 5 RO1 AI22186 from the National Institutes of Health. The Government has certain rights to this invention.

FIELD OF THE INVENTION

10

The present invention relates to recombinant DNA technology, and in particular to introducing and expressing foreign DNA in a eukaryotic cell.

BACKGROUND OF THE INVENTION

15

20

The Alphavirus genus includes a variety of viruses all of which are members of the Togaviridae family. The alphaviruses include Eastern Equine Encephalitis virus (EEE), Venezuelan Equine Encephalitis virus (VEE), Everglades virus, Mucambo virus, Pixuna virus, Western Equine Encephalitis virus (WEE), Sindbis virus, South African Arbovirus No. 86 (S.A.AR 86), Girdwood S.A. virus, Ockelbo virus, Semliki Forest virus, Middelburg virus, Chikungunya virus, O'Nyong-Nyong virus, Ross River virus, Barmah Forest virus, Getah virus, Sagiya virus, Bebaru virus, Mayaro virus, Una virus, Aura virus, Whataroa virus, Babanki virus, Kyzytagach virus, Highlands J virus, Fort Morgan virus, Ndumu virus, and Buggy Creek virus.

-2-

The alphavirus genome is a single-stranded, messenger-sense RNA, modified at the 5'-end with a methylated cap, and at the 3'-end with a variable-length poly (A) tract. The viral genome is divided into two regions: the first encodes the nonstructural or replicase proteins (nsP1-nsP4) and the second encodes the viral structural proteins. Strauss and Strauss, *Microbiological Rev.* 58, 491-562, 494 (1994). Structural subunits consisting of a single viral protein, C, associate with themselves and with the RNA genome in an icosahedral nucleocapsid. In the virion, the capsid is surrounded by a lipid envelope covered with a regular array of transmembranal protein spikes, each of which consists of a heterodimeric complex of two glycoproteins, E1 and E2. See Paredes et al., *Proc. Natl. Acad. Sci. USA* 90, 9095-99 (1993); Paredes et al., *Virology* 187, 324-32 (1993); Pedersen et al., *J. Virol.* 14:40 (1974).

Sindbis virus, the prototype member of the alphavirus genus of the family *Togaviridae*, and viruses related to Sindbis are broadly distributed throughout Africa, Europe, Asia, the Indian subcontinent, and Australia, based on serological surveys of humans, domestic animals and wild birds. Kokernot et al., *Trans. R. Soc. Trop. Med. Hyg.* 59, 553-62 (1965); Redaksie, *S. Afr. Med. J.* 42, 197 (1968); Adekolu-John and Fagbami, *Trans. R. Soc. Trop. Med. Hyg.* 77, 149-51 (1983); Darwish et al., *Trans. R. Soc. Trop. Med. Hyg.* 77, 442-45 (1983); Lundström et al., *Epidemiol. Infect.* 106, 567-74 (1991); Morrill et al., *J. Trop. Med. Hyg.* 94, 166-68 (1991). The first isolate of Sindbis virus (strain AR339) was recovered from a pool of *Culex* sp. mosquitoes collected in Sindbis, Egypt in 1953 (Taylor et al., *Am. J. Trop. Med. Hyg.* 4, 844-62 (1955)), and is the most extensively studied representative of this group. Other members of the Sindbis group of alphaviruses include South African Arbovirus No. 86, Ockelbo82, and Girdwood S.A. These viruses are not strains of the Sindbis virus; they are related to Sindbis AR339, but they are more closely related to each other based on nucleotide sequence and serological comparisons. Lundström et al., *J. Wildl. Dis.* 29, 189-95 (1993); Simpson et al., *Virology* 222, 464-69 (1996). Ockelbo82, S.A.AR86 and Girdwood S.A. are all associated with human disease, whereas Sindbis is not. The clinical symptoms of human infection with Ockelbo82,

-3-

S.A.AR86, or Girdwood S.A. are a febrile illness, general malaise, macropapular rash, and joint pain that occasionally progresses to a polyarthralgia sometimes lasting from a few months to a few years.

5 The study of these viruses has led to the development of beneficial techniques for vaccinating against the alphavirus diseases, and other diseases through the use of alphavirus vectors for the introduction of foreign DNA. *See* United States Patent No. 5,185,440 to Davis et al., and PCT Publication WO 92/10578. It is intended that all United States patent references be incorporated in their entirety by reference.

10 It is well known that live, attenuated viral vaccines are among the most successful means of controlling viral disease. However, for some virus pathogens, immunization with a live virus strain may be either impractical or unsafe. One alternative strategy is the insertion of sequences encoding immunizing antigens of such agents into a vaccine strain of another virus. One such system
15 utilizing a live VEE vector is described in United States Patent No. 5,505,947 to Johnston et al.

Sindbis virus vaccines have been employed as viral carriers in virus constructs which express genes encoding immunizing antigens for other viruses. *See* United States Patent No. 5,217,879 to Huang et al. Huang et al. describes
20 Sindbis infectious viral vectors. However, the reference does not describe the cDNA sequence of Girdwood S.A. and TR339, nor clones or viral vectors produced therefrom.

Another such system is described by Hahn et al., *Proc. Natl. Acad. Sci. USA* 89:2679 (1992), wherein Sindbis virus constructs which express a
25 truncated form of the influenza hemagglutinin protein are described. The constructs are used to study antigen processing and presentation *in vitro* and in mice. Although no infectious challenge dose is tested, it is also suggested that

-4-

such constructs might be used to produce protective B- and T-cell mediated immunity.

London et al., *Proc. Natl. Acad. Sci. USA* 89, 207-11 (1992), disclose a method of producing an immune response in mice against a lethal Rift Valley Fever (RVF) virus by infecting the mice with an infectious Sindbis virus containing an RVF epitope. London does not disclose using Girdwood S.A. or TR339 to induce an immune response in animals.

Viral carriers can also be used to introduce and express foreign DNA in eukaryotic cells. One goal of such techniques is to employ vectors that target expression to particular cells and/or tissues. A current approach has been to remove target cells from the body, culture them *ex vivo*, infect them with an expression vector, and then reintroduce them into the patient.

PCT Publication No. WO 92/10578 to Garoff and Liljeström provide a system for introducing and expressing foreign proteins in animal cells using alphaviruses. This reference discloses the use of Semliki Forest virus to introduce and express foreign proteins in animal cells. The use of Girdwood S.A. or TR339 is not discussed. Furthermore, this reference does not provide a method of targeting and introducing foreign DNA into specific cell or tissue types.

Accordingly, there remains a need in the art for full-length cDNA clones of positive-strand RNA viruses, such as Girdwood S.A and TR339. In addition, there is an ongoing need in the art for improved vaccination strategies. Finally, there remains a need in the art for improved methods and nucleic acid sequences for delivering foreign DNA to target cells.

SUMMARY OF THE INVENTION

A first aspect of the present invention is a method of introducing and expressing heterologous RNA in bone marrow cells, comprising: (a) providing

-5-

a recombinant alphavirus, the alphavirus containing a heterologous RNA segment, the heterologous RNA segment comprising a promoter operable in bone marrow cells operatively associated with a heterologous RNA to be expressed in bone marrow cells; and then (b) contacting the recombinant alphavirus to the bone marrow cells so that the heterologous RNA segment is introduced and expressed therein.

As a second aspect, the present invention provides a helper cell for expressing an infectious, propagation defective, Girdwood S.A. virus particle, comprising, in a Girdwood S.A.-permissive cell: (a) a first helper RNA encoding (i) at least one Girdwood S.A. structural protein, and (ii) not encoding at least one other Girdwood S.A. structural protein; and (b) a second helper RNA separate from the first helper RNA, the second helper RNA (i) not encoding the at least one Girdwood S.A. structural protein encoded by the first helper RNA, and (ii) encoding the at least one other Girdwood S.A. structural protein not encoded by the first helper RNA, and with all of the Girdwood S.A. structural proteins encoded by the first and second helper RNAs assembling together into Girdwood S.A. particles in the cell containing the replicon RNA; and wherein the Girdwood S.A. packaging segment is deleted from at least the first helper RNA.

A third aspect of the present invention is a method of making infectious, propagation defective, Girdwood S.A. virus particles, comprising: transfecting a Girdwood S.A.-permissive cell with a propagation defective replicon RNA, the replicon RNA including the Girdwood S.A. packaging segment and an inserted heterologous RNA; producing the Girdwood S.A. virus particles in the transfected cell; and then collecting the Girdwood S.A. virus particles from the cell. Also disclosed are infectious Girdwood S.A. RNAs, cDNAs encoding the same, infectious Girdwood S.A. virus particles, and pharmaceutical formulations thereof.

As a fourth aspect, the present invention provides a helper cell for expressing an infectious, propagation defective, TR339 virus particle, comprising,

in a TR339-permissive cell: (a) a first helper RNA encoding (i) at least one TR339 structural protein, and (ii) not encoding at least one other TR339 structural protein; and (b) a second helper RNA separate from the first helper RNA, the second helper RNA (i) not encoding the at least one TR339 structural protein encoded by the first helper RNA, and (ii) encoding the at least one other TR339 structural protein not encoded by the first helper RNA, and with all of the TR339 structural proteins encoded by the first and second helper RNAs assembling together into TR339 particles in the cell containing the replicon RNA; and wherein the TR339 packaging segment is deleted from at least the first helper RNA.

A fifth aspect of the present invention is a method of making infectious, propagation defective, TR339 virus particles, comprising: transfecting a TR339-permissive cell with a propagation defective replicon RNA, the replicon RNA including the TR339 packaging segment and an inserted heterologous RNA; producing the TR339 virus particles in the transfected cell; and then collecting the TR339 virus particles from the cell. Also disclosed are infectious TR339 RNAs, cDNAs encoding the same, infectious TR339 virus particles, and pharmaceutical formulations thereof.

As a sixth aspect, the present invention provides a recombinant DNA comprising a cDNA coding for an infectious Girdwood S.A. virus RNA transcript, and a heterologous promoter positioned upstream from the cDNA and operatively associated therewith. The present invention also provides infectious RNA transcripts encoded by the above-mentioned cDNA and infectious viral particles containing the infectious RNA transcripts.

As a seventh aspect, the present invention provides a recombinant DNA comprising a cDNA coding for a Sindbis strain TR339 RNA transcript, and a heterologous promoter positioned upstream from the cDNA and operatively associated therewith. The present invention also provides infectious RNA transcripts encoded by the above-mentioned cDNA and infectious viral particles containing the infectious RNA transcripts.

The foregoing and other aspects of the present invention are described in the detailed description set forth below.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Figure 1 presents the cDNA sequence (SEQ ID NO:1) of S.A.AR86. The RNA sequence of the 5' 40 nucleotides was obtained by direct sequencing of the genomic RNA. The rest of the genome was sequenced by RT-PCR of fragments amplified from virion RNA. Nucleotides 1 through 59 represent the 5' UTR, the non-structural polyprotein is encoded by nucleotides 60 through 7559 (nsP1--nt60 through nt1679; nsP2--nt1680 through nt4099; nsP3--
10 nt4100 through nt5729; nsP4--nt5730 through nt7559), the structural polyprotein is encoded by nucleotides 7608 through 11342 (capsid--nt7608 through nt8399; E3--nt8400 through nt8591; E2--nt8592 through nt9860; 6K--nt9861 through nt10025; E1--nt10026 through nt11342), and the 3' UTR is represented by nucleotides 11346 through 11663.

15 Figure 1A shows nucleotides 1 through 3800 of the cDNA sequence of S.A.AR86.

Figure 1B shows nucleotides 3801 through 7900 of the cDNA sequence of S.A.AR86.

20 Figure 1C shows nucleotides 7901 through 11663 of the cDNA sequence of S.A.AR86.

Figure 2 presents the putative amino acid sequences of the S.A.AR86 polyproteins (SEQ ID NO:2 and SEQ ID NO:3). The amino acids were derived from the S.A.AR86 cDNA sequence given in Figure 1 (SEQ ID NO:1).

-8-

Figure 2A shows the amino acid sequence of the non-structural polyprotein of S.A.AR86 (SEQ ID NO:2).

Figure 2B shows the amino acid sequence of the structural polyprotein of S.A.AR86 (SEQ ID NO:3).

5 Figure 3 presents the cDNA sequence (SEQ ID NO:4) of Girdwood S.A. The RNA sequence of the 5' 40 nucleotides was obtained by direct sequencing of the genomic RNA. The rest of the genome sequence was obtained by sequencing of fragments amplified by RT-PCR from virion RNA. An "N" in the sequence indicates that the identity of the nucleotide at that position is
10 unknown. Nucleotides 1 through 59 represent the 5' UTR, the non-structural polyprotein is encoded by nucleotides 60 through 7613 (nsP1--nt60 through nt1679; nsP2--nt1680 through nt4099; nsP3--nt4100 through nt5762 or nt5783; nsP4--nt5784 through nt7613), the structural polyprotein is encoded by nucleotides
15 7662 through 11396 (capsid--nt7662 through nt8453; E3--nt8454 through nt8645; E2--nt8646 through nt9914, 6K--9915 through nt10079; E1--nt10080 through nt11396), and the 3' UTR is represented by nucleotides 11400 through 11717. There is an opal termination codon at nucleotides 5763 through 5765.

Figure 3A shows nucleotides 1 through 3800 of the cDNA sequence of Girdwood S.A.

20 Figure 3B shows nucleotides 3801 through 7900 of the cDNA sequence of Girdwood S.A.

Figure 3C shows nucleotides 7901 through 11717 of the cDNA sequence of Girdwood S.A.

25 Figure 4 illustrates the putative amino acid sequences of the Girdwood S.A. polyproteins (SEQ ID NO:5 and SEQ ID NO:6). The amino

-9-

acids were derived from the Girdwood S.A. cDNA sequence given in Figure 3 (SEQ ID NO:4).

5 Figure 4A shows the amino acid sequence of the non-structural polyprotein of Girdwood S.A. The sequence terminates at the opal termination codon. The complete amino acid sequence is presented in SEQ ID NO:5.

 Figure 4B shows the amino acid sequence of the structural polyprotein of Girdwood S.A. (SEQ ID NO:6).

 Figure 5 illustrates the nucleotide sequence (SEQ ID NO:7) of clone pS55, a cDNA clone of the S.A.AR86 genomic RNA.

10 Figure 5A shows nucleotides 1 through 6720 of the cDNA sequence of pS55.

 Figure 5B shows nucleotides 6721 through 11663 of the cDNA sequence of pS55.

15 Figure 6 presents the cDNA sequence (SEQ ID NO:8) of clone pTR339. The TR339 virus is derived from this clone. Nucleotides 1 through 59 represent the 5' UTR, the non-structural polyprotein is encoded by nucleotides 60 through 7598 (nsP1--nt60 through nt1679; nsP2--nt1680 through nt4099; nsP3--nt4100 through nt5747 or 5768; nsP4--nt5769 through nt7598), the structural polyprotein is encoded by nucleotides 7647 through 11381 (capsid--nt7647 through nt8438; E3--nt8439 through nt8630; E2--nt8631 through nt9899; 6K--nt9900 through nt10064; E1--nt10065 through nt11381), and the 3' UTR is represented by nucleotides 11382 through 11703. There is an opal termination codon at nucleotides 5748 through 5750.

20 Figure 6A shows nucleotides 1 through 6720 of the cDNA sequence of pTR339.

-10-

Figure 6B shows nucleotides 6721 through 11703 of the cDNA sequence of pTR339.

DETAILED DESCRIPTION OF THE INVENTION

5 The production and use of recombinant DNA, vectors, transformed host cells, selectable markers, proteins, and protein fragments by genetic engineering are well-known to those skilled in the art. *See, e.g.*, United States Patent No. 4,761,371 to Bell et al. at Col. 6 line 3 to Col. 9 line 65; United States Patent No. 4,877, 729 to Clark et al. at Col. 4 line 38 to Col. 7 line 6; United States Patent No. 4,912,038 to Schilling at Col 3 line 26 to Col 14 line 12; and
10 United States Patent No. 4,879,224 to Wallner at Col. 6 line 8 to Col. 8 line 59.

The term "alphavirus" has its conventional meaning in the art, and includes the various species of alphaviruses such as Eastern Equine Encephalitis virus (EEE), Venezuelan Equine Encephalitis virus (VEE), Everglades virus, Mucambo virus, Pixuna virus, Western Encephalitis virus (WEE), Sindbis virus,
15 South African Arbovirus No. 86, Girdwood S.A. virus, Ockelbo virus, Semliki Forest virus, Middelburg virus, Chikungunya virus, O'Nyong-Nyong virus, Ross River virus, Barmah Forest virus, Getah virus, Sagiyama virus, Bebaru virus, Mayaro virus, Una virus, Aura virus, Whataroa virus, Babanki virus, Kyzlagach virus, Highlands J virus, Fort Morgan virus, Ndumu virus, Buggy Creek virus,
20 and any other virus classified by the International Committee on Taxonomy of Viruses (ICTV) as an alphavirus. The preferred alphaviruses for use in the present invention include Sindbis virus strains (*e.g.*, TR339), Girdwood S.A., S.A.AR86, and Ockelbo82.

An "Old World alphavirus" is a virus that is primarily distributed
25 throughout the Old World. Alternately stated, an Old World alphavirus is a virus that is primarily distributed throughout Africa, Asia, Australia and New Zealand, or Europe. Exemplary Old World viruses include SF group alphaviruses and SIN group alphaviruses. SF group alphaviruses include Semliki Forest virus, Middelburg virus, Chikungunya virus, O'Nyong-Nyong virus, Ross River virus,

-11-

Barmah Forest virus, Getah virus, Sagiyama virus, Bebaru virus, Mayaro virus, and Una virus. SIN group alphaviruses include Sindbis virus, South African Arbovirus No. 86, Ockelbo virus, Girdwood S.A. virus, Aura virus, Whataroa virus, Babanki virus, and Kyzylagach virus.

5 Acceptable alphaviruses include those containing attenuating mutations. The phrases "attenuating mutation" and "attenuating amino acid," as used herein, mean a nucleotide sequence containing a mutation, or an amino acid encoded by a nucleotide sequence containing a mutation, which mutation results in a decreased probability of causing disease in its host (*i.e.*, a loss of virulence),
10 in accordance with standard terminology in the art, whether the mutation be a substitution mutation or an in-frame deletion mutation. *See, e.g.*, B. DAVIS ET AL., MICROBIOLOGY 132 (3d ed. 1980). The phrase "attenuating mutation" excludes mutations or combinations of mutations which would be lethal to the virus.

15 Appropriate attenuating mutations will be dependent upon the alphavirus used. Suitable attenuating mutations within the alphavirus genome will be known to those skilled in the art. Exemplary attenuating mutations include, but are not limited to, those described in United States Patent No. 5,505,947 to Johnston et al., copending United States application 08/448,630 to Johnston et al.,
20 and copending United States application 08/446,932 to Johnston et al. It is intended that all United States patent references be incorporated in their entirety by reference.

25 Attenuating mutations may be introduced into the RNA by performing site-directed mutagenesis on the cDNA which encodes the RNA, in accordance with known procedures. *See*, Kunkel, *Proc. Natl. Acad. Sci. USA* 82, 488 (1985), the disclosure of which is incorporated herein by reference in its entirety. Alternatively, mutations may be introduced into the RNA by replacement of homologous restriction fragments in the cDNA which encodes for the RNA, in accordance with known procedures.

I. Methods for Introducing and Expressing Heterologous RNA in Bone Marrow Cells.

The present invention provides methods of using a recombinant alphavirus to introduce and express a heterologous RNA in bone marrow cells. Such methods are useful as vaccination strategies when the heterologous RNA encodes an immunogenic protein or peptide. Alternatively, such methods are useful in introducing and expressing in bone marrow cells an RNA which encodes a desirable protein or peptide, for example, a therapeutic protein or peptide.

The present invention is carried out using a recombinant alphavirus to introduce a heterologous RNA into bone marrow cells. Any alphavirus that targets and infects bone marrow cells is suitable. Preferred alphaviruses include Old World alphaviruses, more preferably SF group alphaviruses and SIN group alphaviruses, more preferably Sindbis virus strains (*e.g.*, TR339), S.A.AR86 virus, Girdwood S.A. virus, and Ockelbo virus. In a more preferred embodiment, the alphavirus contains one or more attenuating mutations, as described hereinabove.

Two types of recombinant virus vector are contemplated in carrying out the present invention. In one embodiment employing "double promoter vectors," the heterologous RNA is inserted into a replication and propagation competent virus. Double promoter vectors are described in United States Patent No. 5,505,947 to Johnston et al. With this type of viral vector, it is preferable that heterologous RNA sequences of less than 3 kilobases are inserted into the viral vector, more preferably those less than 2 kilobases, and more preferably still those less than 1 kilobase. In an alternate embodiment, propagation-defective "replicon vectors," as described in copending United States application 08/448,630 to Johnston et al., will be used. One advantage of replicon viral vectors is that larger RNA inserts, up to approximately 4-5 kilobases in length can be utilized. Double promoter vectors and replicon vectors are described in more detail hereinbelow.

-13-

The recombinant alphaviruses of the claimed method target the heterologous RNA to bone marrow cells, where it expresses the encoded protein or peptide. Heterologous RNA can be introduced and expressed in any cell type found in the bone marrow. Bone marrow cells that may be targeted by the recombinant alphaviruses of the present invention include, but are not limited to, polymorphonuclear cells, hemopoietic stem cells (including megakaryocyte colony forming units (CFU-M), spleen colony forming units (CFU-S), erythroid colony forming units (CFU-E), erythroid burst forming units (BFU-E), and colony forming units in culture (CFU-C), erythrocytes, macrophages (including reticular cells), monocytes, granulocytes, megakaryocytes, lymphocytes, fibroblasts, osteoprogenitor cells, osteoblasts, osteoclasts, marrow stromal cells, chondrocytes and other cells of synovial joints. Preferably, marrow cells within the endosteum are targeted, more preferably osteoblasts. Also preferred are methods in which cells in the endosteum of synovial joints (*e.g.*, hip and knee joints) are targeted.

By targeting to the cells of the bone marrow, it is meant that the primary site in which the virus will be localized *in vivo* is the cells of the bone marrow. Alternately stated, the alphaviruses of the present invention target bone marrow cells, such that titers in bone marrow two days after infection are greater than 100 PFU/g crushed bone, preferably greater than 200 PFU/g crushed bone, more preferably greater than 300 PFU/g crushed bone, and more preferably still greater than 500 PFU/g crushed bone. Virus may be detected occasionally in other cell or tissue types, but only sporadically and usually at low levels. Virus localization in the bone marrow can be demonstrated by any suitable technique known in the art, such as *in situ* hybridization.

Bone marrow cells are long-lived and harbor infectious alphaviruses for a prolonged period of time, as demonstrated in the Examples below. These characteristics of bone marrow cells render the present invention useful not only for the purpose of supplying a desired protein or peptide to skeletal tissue, but also for expressing proteins or peptides *in vivo* that are needed by other cell or tissue types.

The present invention can be carried out *in vivo* or with cultured bone marrow cells *in vitro*. Bone marrow cell cultures include primary cultures

-14-

of bone marrow cells, serially-passaged cultures of bone marrow cells, and cultures of immortalized bone marrow cell lines. Bone marrow cells may be cultured by any suitable means known in the art.

5 The recombinant alphaviruses of the present invention carry a heterologous RNA segment. The heterologous RNA segment encodes a promoter and an inserted heterologous RNA. The inserted heterologous RNA may encode any protein or a peptide which is desirably expressed by the host bone marrow cells. Suitable heterologous RNA may be of prokaryotic (*e.g.*, RNA encoding the *Botulinus* toxin C), or eukaryotic (*e.g.*, RNA encoding malaria *Plasmodium* protein cs1) origin. Illustrative proteins and peptides encoded by the heterologous
10 RNAs of the present invention include hormones, growth factors, interleukins, cytokines, chemokines, enzymes, and ribozymes. Alternately, the heterologous RNAs encode any therapeutic protein or peptide. As a further alternative, the heterologous RNAs of the present invention encode any immunogenic protein or
15 peptide.

An immunogenic protein or peptide, or "immunogen," may be any protein or peptide suitable for protecting the subject against a disease, including but not limited to microbial, bacterial, protozoal, parasitic, and viral diseases. For example, the immunogen may be an orthomyxovirus immunogen (*e.g.*, an
20 influenza virus immunogen, such as the influenza virus hemagglutinin (HA) surface protein or the influenza virus nucleoprotein gene, or an equine influenza virus immunogen), or a lentivirus immunogen (*e.g.*, an equine infectious anemia virus immunogen, a Simian Immunodeficiency Virus (SIV) immunogen, or a Human Immunodeficiency Virus (HIV) immunogen, such as the HIV envelope
25 GP160 protein and the HIV matrix/capsid proteins). The immunogen may also be an arenavirus immunogen (*e.g.*, Lassa fever virus immunogen, such as the Lassa fever virus nucleocapsid protein gene and the Lassa fever envelope glycoprotein gene), a poxvirus immunogen (*e.g.*, vaccinia), a flavivirus immunogen (*e.g.*, a yellow fever virus immunogen or a Japanese encephalitis virus immunogen), a
30 filovirus immunogen (*e.g.*, an Ebola virus immunogen, or a Marburg virus

immunogen), a bunyavirus immunogen (*e.g.*, RVFV, CCHF, and SFS viruses),
or a coronavirus immunogen (*e.g.*, an infectious human coronavirus immunogen,
such as the human coronavirus envelope glycoprotein gene, or a transmissible
gastroenteritis virus immunogen for pigs, or an infectious bronchitis virus
immunogen for chickens).

Alternatively, the present invention can be used to express
heterologous RNAs encoding antisense oligonucleotides. In general, "antisense"
refers to the use of small, synthetic oligonucleotides to inhibit gene expression by
inhibiting the function of the target mRNA containing the complementary
sequence. Milligan, J.F. et al., *J. Med. Chem.* 36(14), 1923-1937 (1993). Gene
expression is inhibited through hybridization to coding (sense) sequences in a
specific mRNA target by hydrogen bonding according to Watson-Crick base
pairing rules. The mechanism of antisense inhibition is that the exogenously
applied oligonucleotides decrease the mRNA and protein levels of the target gene.
Milligan, J.F. et al., *J. Med. Chem.* 36(14), 1923-1937 (1993). *See also* Helene,
C. and Toulme, J., *Biochim. Biophys. Acta* 1049, 99-125 (1990); Cohen, J.S.,
Ed., OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE
EXPRESSION, CRC Press:Boca Raton, FL (1987).

Antisense oligonucleotides may be of any suitable length, depending
on the particular target being bound. The only limits on the length of the antisense
oligonucleotide is the capacity of the virus for inserted heterologous RNA.
Antisense oligonucleotides may be complementary to the entire mRNA transcript
of the target gene or only a portion thereof. Preferably the antisense
oligonucleotide is directed to an mRNA region containing a junction between
intron and exon. Where the antisense oligonucleotide is directed to an intron/exon
junction, it may either entirely overlie the junction or may be sufficiently close to
the junction to inhibit splicing out of the intervening exon during processing of
precursor mRNA to mature mRNA (*e.g.*, with the 3' or 5' terminus of the
antisense oligonucleotide being positioned within about, for example, 10, 5, 3 or

-16-

2 nucleotides of the intron/exon junction). Also preferred are antisense oligonucleotides which overlap the initiation codon.

When practicing the present invention, the antisense oligonucleotides administered may be related in origin to the species to which it is administered.

5 When treating humans, human antisense may be used if desired.

Promoters for use in carrying out the present invention are operable in bone marrow cells. An operable promoter in bone marrow cells is a promoter that is recognized by and functions in bone marrow cells. Promoters for use with the present invention must also be operatively associated with the heterologous
10 RNA to be expressed in the bone marrow. A promoter is operably linked to a heterologous RNA if it controls the transcription of the heterologous RNA, where the heterologous RNA comprises a coding sequence. Suitable promoters are well known in the art. The Sindbis 26S promoter is preferred when the alphavirus is a strain of Sindbis virus. Additional preferred promoters beyond the Sindbis 26S
15 promoter include the Girdwood S.A. 26S promoter when the alphavirus is Girdwood S.A., the S.A.AR86 26S promoter when the alphavirus is S.A.AR86, and any other promoter sequence recognized by alphavirus polymerases. Alphavirus promoter sequences containing mutations which alter the activity level of the promoter (in relation to the activity level of the wild-type) are also suitable
20 in the practice of the present invention. Such mutant promoter sequences are described in Raju and Huang, *J. Virol.* 65, 2501-2510 (1991), the disclosure of which is incorporated in its entirety by reference.

The heterologous RNA is introduced into the bone marrow cells by contacting the recombinant alphavirus carrying the heterologous RNA segment to
25 the bone marrow cells. By contacting, it is meant bringing the recombinant alphavirus and the bone marrow cells in physical proximity. The contacting step can be performed *in vitro* or *in vivo*. *In vitro* contacting can be carried out with cultures of immortalized or non-immortalized bone marrow cells. In one particular embodiment, bone marrow cells can be removed from a subject, cultured *in vitro*,

infected with the vector, and then introduced back into the subject. Contacting is performed *in vivo* when the recombinant alphavirus is administered to a subject. Pharmaceutical formulations of recombinant alphavirus can be administered to a subject parenterally (*e.g.*, subcutaneous, intracerebral, intradermal, intramuscular, intravenous and intraarticular) administration. Alternatively, pharmaceutical formulations of the present invention may be suitable for administration to the mucus membranes of a subject (*e.g.*, intranasal administration, by use of a dropper, swab, or inhaler). Methods of preparing infectious virus particles and pharmaceutical formulations thereof are discussed in more detail hereinbelow.

By "introducing" the heterologous RNA segment into the bone marrow cells it is meant infecting the bone marrow cells with recombinant alphavirus containing the heterologous RNA, such that the viral vector carrying the heterologous RNA enters the bone marrow cells and can be expressed therein. As used with respect to the present invention, when the heterologous RNA is "expressed," it is meant that the heterologous RNA is transcribed. In particular embodiments of the invention in which it is desired to produce a protein or peptide, expression further includes the steps of post-transcriptional processing and translation of the mRNA transcribed from the heterologous RNA. In contrast, where the heterologous RNA encodes an antisense oligonucleotide, expression need not include post-transcriptional processing and translation. With respect to embodiments in which the heterologous RNA encodes an immunogenic protein or a protein being administered for therapeutic purposes, expression may also include the further step of post-translational processing to produce an immunogenic or therapeutically-active protein.

The present invention also provides infectious RNAs, as described hereinabove, and cDNAs encoding the same. Preferably the infectious RNAs and cDNAs are derived from the S.A.AR86, Girdwood S.A., TR339, or Ockelbo viruses. The cDNA clones can be generated by any of a variety of suitable methods known to those skilled in the art. A preferred method is the method set

-18-

forth in United States Patent No. 5,185,440 to Davis et al., the disclosure of which is incorporated in its entirety by reference, and Gubler et al., *Gene* 25:263 (1983).

RNA is preferably synthesized from the DNA sequence *in vitro* using purified RNA polymerase in the presence of ribonucleotide triphosphates and cap analogs in accordance with conventional techniques. However, the RNA may also be synthesized intracellularly after introduction of the cDNA.

A. Double Promoter Vectors.

In one embodiment of the invention, double promoter vectors are used to introduce the heterologous RNA into the target bone marrow cells. A double promoter virus vector is a replication and propagation competent virus. Double promoter vectors are described in United States Patent No. 5,505,947 to Johnston et al., the disclosure of which is incorporated in its entirety by reference. Preferred alphaviruses for constructing the double promoter vectors are S.A.AR86, Girdwood S.A., TR339 and Ockelbo viruses. More preferably, the double promoter vector contains one or more attenuating mutations. Attenuating mutations are described in more detail hereinabove.

The double promoter vector is constructed so as to contain a second subgenomic promoter (*i.e.*, 26S promoter) inserted 3' to the virus RNA encoding the structural proteins. The heterologous RNA is inserted between the second subgenomic promoter, so as to be operatively associated therewith, and the 3' UTR of the virus genome. Heterologous RNA sequences of less than 3 kilobases, more preferably those less than 2 kilobases, and more preferably still those less than 1 kilobase, can be inserted into the double promoter vector. In a preferred embodiment of the invention, the double promoter vector is derived from Girdwood S.A., and the second subgenomic promoter is a duplicate of the Girdwood S.A. subgenomic promoter. In an alternate preferred embodiment, the double promoter vector is derived from TR339, and the second subgenomic promoter is a duplicate of the TR339 subgenomic promoter.

B. Replicon Vectors.

Replicon vectors, which are propagation-defective virus vectors can also be used to carry out the present invention. Replicon vectors are described in more detail in copending United States Application 08/448,630 to Johnston et al., the disclosure of which is incorporated in its entirety by reference. Preferred alphaviruses for constructing the replicon vectors are S.A.AR86, Girdwood S.A., TR339, and Ockelbo.

In general, in the replicon system, a foreign gene to be expressed is inserted in place of at least one of the viral structural protein genes in a transcription plasmid containing an otherwise full-length cDNA copy of the alphavirus genome RNA. RNA transcribed from this plasmid contains an intact copy of the viral nonstructural genes which are responsible for RNA replication and transcription. Thus, if the transcribed RNA is transfected into susceptible cells, it will be replicated and translated to give the nonstructural proteins. These proteins will transcribe the transfected RNA to give high levels of subgenomic mRNA, which will then be translated to produce high levels of the foreign protein. The autonomously replicating RNA (*i.e.*, replicon) can only be packaged into virus particles if the alphavirus structural protein genes are provided on one or more "helper" RNAs, which are cotransfected into cells along with the replicon RNA. The helper RNAs do not contain the viral nonstructural genes for replication, but these functions are provided *in trans* by the replicon RNA. Similarly, the transcriptase functions translated from the replicon RNA transcribe the structural protein genes on the helper RNA, resulting in the synthesis of viral structural proteins and packaging of the replicon into virus-like particles. As the packaging or encapsidation signal for alphavirus RNAs is located within the nonstructural genes, the absence of these sequences in the helper RNAs precludes their incorporation into virus particles.

Alphavirus-permissive cells employed in the methods of the present invention are cells which, upon transfection with the viral RNA transcript, are capable of producing viral particles. Preferred alphavirus-permissive cells are

-20-

TR339-permissive cells, Girdwood S.A.-permissive cells, S.A.AR86-permissive cells, and Ockelbo-permissive cells. Alphaviruses have a broad host range. Examples of suitable host cells include, but are not limited to Vero cells, baby hamster kidney (BHK) cells, and chicken embryo fibroblast cells.

5 The phrase "structural protein" as used herein refers to the encoded proteins which are required for encapsidation (*e.g.*, packaging) of the RNA replicon, and include the capsid protein, E1 glycoprotein, and E2 glycoprotein. As described hereinabove, the structural proteins of the alphavirus are distributed among one or more helper RNAs (*i.e.*, a first helper RNA and a second helper RNA). In addition, one or
10 more structural proteins may be located on the same RNA molecule as the replicon RNA, provided that at least one structural protein is deleted from the replicon RNA such that the resulting alphavirus particle is propagation defective. As used herein, the terms "deleted" or "deletion" mean either total deletion of the specified segment or the deletion of a sufficient portion of the specified segment to render the segment inoperative or
15 nonfunctional, in accordance with standard usage. *See, e.g.*, U.S. Patent No. 4,650,764 to Temin et al. The term "propagation defective" as used herein, means that the replicon RNA cannot be encapsidated in the host cell in the absence of the helper RNA. The resulting alphavirus replicon particles are propagation defective inasmuch as the replicon RNA in these particles does not include all of the alphavirus structural proteins required
20 for encapsidation, at least one of the required structural proteins being deleted therefrom, such that the replicon RNA initiates only an abortive infection; no new viral particles are produced, and there is no spread of the infection to other cells.

The helper cell for expressing the infectious, propagation defective alphavirus particle comprises a set of RNAs, as described above. The set of RNAs principally
25 include a first helper RNA and a second helper RNA. The first helper RNA includes RNA encoding at least one alphavirus structural protein but does not encode all alphavirus structural proteins. In other words, the first helper RNA does not encode at least one alphavirus structural protein; the at least one non-coded alphavirus structural protein being deleted from the first helper RNA.

-21-

In one embodiment, the first helper RNA includes RNA encoding the alphavirus E1 glycoprotein, with the alphavirus capsid protein and the alphavirus E2 glycoprotein being deleted from the first helper RNA. In another embodiment, the first helper RNA includes RNA encoding the alphavirus E2 glycoprotein, with the alphavirus capsid protein and the alphavirus E1 glycoprotein being deleted from the first helper RNA. In a third, preferred embodiment, the first helper RNA includes RNA encoding the alphavirus E1 glycoprotein and the alphavirus E2 glycoprotein, with the alphavirus capsid protein being deleted from the first helper RNA.

The second helper RNA includes RNA encoding at least one alphavirus structural protein which is different from the at least one structural protein encoded by the first helper RNA. Thus, the second helper RNA encodes at least one alphavirus structural protein which is not encoded by the first helper RNA. The second helper RNA does not encode the at least one alphavirus structural protein which is encoded by the first helper RNA, thus the first and second helper RNAs do not encode duplicate structural proteins. In the embodiment wherein the first helper RNA includes RNA encoding only the alphavirus E1 glycoprotein, the second helper RNA may include RNA encoding one or both of the alphavirus capsid protein and the alphavirus E2 glycoprotein which are deleted from the first helper RNA. In the embodiment wherein, the first helper RNA includes RNA encoding only the alphavirus E2 glycoprotein, the second helper RNA may include RNA encoding one or both of the alphavirus capsid protein and the alphavirus E1 glycoprotein which are deleted from the first helper RNA. In the embodiment wherein the first helper RNA includes RNA encoding both the alphavirus E1 glycoprotein and the alphavirus E2 glycoprotein, the second helper RNA may include RNA encoding the alphavirus capsid protein which is deleted from the first helper RNA.

In one embodiment, the packaging segment (RNA comprising the encapsidation or packaging signal) is deleted from at least the first helper RNA.

-22-

In a preferred embodiment, the packaging segment is deleted from both the first helper RNA and the second helper RNA.

In the preferred embodiment wherein the packaging segment is deleted from both the first helper RNA and the second helper RNA, the helper cell is co-transfected with a replicon RNA in addition to the first helper RNA and the second helper RNA. The replicon RNA encodes the packaging segment and an inserted heterologous RNA. The inserted heterologous RNA may be RNA encoding a protein or a peptide. In a preferred embodiment, the replicon RNA, the first helper RNA and the second helper RNA are provided on separate molecules such that a first molecule, *i.e.*, the replicon RNA, includes RNA encoding the packaging segment and the inserted heterologous RNA, a second molecule, *i.e.*, the first helper RNA, includes RNA encoding at least one but not all of the required alphavirus structural proteins, and a third molecule, *i.e.*, the second helper RNA, includes RNA encoding at least one but not all of the required alphavirus structural proteins. For example, in one preferred embodiment of the present invention, the helper cell includes a set of RNAs which include (a) a replicon RNA including RNA encoding an alphavirus packaging sequence and an inserted heterologous RNA, (b) a first helper RNA including RNA encoding the alphavirus E1 glycoprotein and the alphavirus E2 glycoprotein, and (c) a second helper RNA including RNA encoding the alphavirus capsid protein so that the alphavirus E1 glycoprotein, the alphavirus E2 glycoprotein and the capsid protein assemble together into alphavirus particles in the host cell.

In an alternate embodiment, the replicon RNA and the first helper RNA are on separate molecules, and the replicon RNA and RNA encoding a structural gene not encoded by the first helper RNA are on another single molecule together, such that a first molecule, *i.e.*, the first helper RNA, including RNA encoding at least one but not all of the required alphavirus structural proteins, and a second molecule, *i.e.*, the replicon RNA, including RNA encoding the packaging segment, the inserted heterologous RNA, and the remaining structural proteins not encoded by the first helper RNA. For example, in one preferred embodiment of

the present invention, the helper cell includes a set of RNAs including (a) a replicon RNA including RNA encoding an alphavirus packaging sequence, an inserted heterologous RNA, and an alphavirus capsid protein, and (b) a first helper RNA including RNA encoding the alphavirus E1 glycoprotein and the alphavirus E2 glycoprotein so that the alphavirus E1 glycoprotein, the alphavirus E2 glycoprotein and the capsid protein assemble together into alphavirus particles in the host cell, with the replicon RNA packaged therein.

In one preferred embodiment of the present invention, the RNA encoding the alphavirus structural proteins, *i.e.*, the capsid, E1 glycoprotein and E2 glycoprotein, contains at least one attenuating mutation, as described hereinabove. Thus, according to this embodiment, at least one of the first helper RNA and the second helper RNA includes at least one attenuating mutation. In a more preferred embodiment, at least one of the first helper RNA and the second helper RNA includes at least two, or multiple, attenuating mutations. The multiple attenuating mutations may be positioned in either the first helper RNA or in the second helper RNA, or they may be distributed randomly with one or more attenuating mutations being positioned in the first helper RNA and one or more attenuating mutations positioned in the second helper RNA. Alternatively, when the replicon RNA and the RNA encoding the structural proteins not encoded by the first helper RNA are located on the same molecule, an attenuating mutation may be positioned in the RNA which codes for the structural protein not encoded by the first helper RNA. The attenuating mutations may also be located within the RNA encoding non-structural proteins (*e.g.*, the replicon RNA).

Preferably, the first helper RNA and the second helper RNA also include a promoter. It is also preferred that the replicon RNA also includes a promoter. Suitable promoters for inclusion in the first helper RNA, second helper RNA and replicon RNA are well known in the art. One preferred promoter is the Girdwood S.A. 26S promoter for use when the alphavirus is Girdwood S.A. Another preferred promoter is the TR339 26S promoter for use when the alphavirus is TR339. Additional promoters beyond the Girdwood S.A. and TR339

promoters include the VEE 26S promoter, the Sindbis 26S promoter, the Semliki Forest 26S promoter, and any other promoter sequence recognized by alphavirus polymerases. Alphavirus promoter sequences containing mutations which alter the activity level of the promoter (in relation to the activity level of the wild-type) are also suitable in the practice of the present invention. Such mutant promoter sequences are described in Raju and Huang, *J. Virol.* 65, 2501-2510 (1991), the disclosure of which is incorporated herein in its entirety. In the system wherein the first helper RNA, the second helper RNA, and the replicon RNA are all on separate molecules, the promoters, if the same promoter is used for all three RNAs, provide a homologous sequence between the three molecules. It is preferred that the selected promoter is operative with the non-structural proteins encoded by the replicon RNA molecule.

In cases where vaccination with two immunogens provides improved protection against disease as compared to vaccination with only a single immunogen, a double-promoter replicon would ensure that both immunogens are produced in the same cell. Such a replicon would be the same as the one described above, except that it would contain two copies of the 26S RNA promoter, each followed by a different multiple cloning site, to allow for the insertion and expression of two different heterologous proteins. Another useful strategy is to insert the IRES sequence from the picornavirus, EMC virus, between the two heterologous genes downstream from the single 26S promoter of the replicon described above, thus leading to expression of two immunogens from the single replicon transcript in the same cell.

C. Uses of the Present Invention.

The alphavirus vectors, RNAs, cDNAs, helper cells, infectious virus particles, and methods of the present invention find use in *in vitro* expression systems, wherein the inserted heterologous RNA encodes a protein or peptide which is desirably produced *in vitro*. The RNAs, cDNAs, helper cells, infectious virus particles, methods, and pharmaceutical formulations of the present invention are additionally useful in a method of administering a protein or peptide to a

subject in need of the protein or peptide, as a method of treatment or otherwise. In this embodiment of the invention, the heterologous RNA encodes the desired protein or peptide, and pharmaceutical formulations of the present invention are administered to a subject in need of the desired protein or peptide. In this manner,
5 the protein or peptide may thus be produced *in vivo* in the subject. The subject may be in need of the protein or peptide because the subject has a deficiency thereof, or because the production of the protein or peptide in the subject may impart some therapeutic effect, as a method of treatment or otherwise.

Alternately, the claimed methods provide a vaccination strategy,
10 wherein the heterologous RNA encodes an immunogenic protein or peptide.

The methods and products of the invention are also useful as antigens and for evoking the production of antibodies in animals such as horses and rabbits, from which the antibodies may be collected and then used in diagnostic assays in accordance with known techniques.

15 A further aspect of the present invention is a method of introducing and expressing antisense oligonucleotides in bone marrow cell cultures to regulate gene expression. Alternately, the claimed method finds use in introducing and expressing a protein or peptide in bone marrow cell cultures.

II. Girdwood S.A. and TR339 Clones.

20 Disclosed hereinbelow are genomic RNA sequences encoding live Girdwood S.A. virus, live S.A.AR86 virus, and live Sindbis strain TR339 virus, cDNAs derived therefrom, infectious RNA transcripts encoded by the cDNAs, infectious viral particles containing the infectious RNA transcripts, and pharmaceutical formulations derived therefrom.

25 The cDNA sequence of Girdwood S.A. is given herein as SEQ ID NO:4. Alternatively, the cDNA may have a sequence which differs from the cDNA of SEQ ID NO:4, but which has the same protein sequence as the cDNA

-26-

given herein as SEQ ID NO:4. Thus, the cDNA may include one or more silent mutations.

5 The phrase "silent mutation" as used herein refers to mutations in the cDNA coding sequence which do not produce mutations in the corresponding protein sequence translated therefrom.

10 Likewise, the cDNA sequence of TR339 is given herein as SEQ ID NO:8. Alternatively, the cDNA may have a sequence which differs from the cDNA of SEQ ID NO:8, but which has the same protein sequence as the cDNA given herein as SEQ ID NO:8. Thus, the cDNA may include one or more silent mutations.

15 The cDNAs encoding infectious Girdwood S.A. and TR339 virus RNA transcripts of the present invention include those homologous to, and having essentially the same biological properties as, the cDNA sequences disclosed herein as SEQ ID NO:4 and SEQ ID NO:8, respectively. Thus, cDNAs that hybridize to cDNAs encoding infectious Girdwood S.A. or TR339 virus RNA transcripts disclosed herein are also an aspect of this invention. Conditions which will permit other cDNAs encoding infectious Girdwood S.A. or TR339 virus transcripts to hybridize to the cDNAs disclosed herein can be determined in accordance with known techniques. For example, hybridization of such sequences may be carried out under conditions of reduced stringency, medium stringency, or even high stringency conditions (*e.g.*, conditions represented by a wash stringency of 35-40% formamide with 5X Denhardt's solution, 0.5% SDS and 1X SSPE at 37°C; conditions represented by a wash stringency of 40-45% formamide with 5X Denhardt's solution, 0.5% SDS, and 1X SSPE at 42°C; and conditions represented by a wash stringency of 50% formamide with 5X Denhardt's solution, 0.5% SDS and 1X SSPE at 42°C, respectively, to cDNA encoding infectious Girdwood S.A. or TR339 virus RNA transcripts disclosed herein in a standard hybridization assay. *See J. SAMBROOK ET AL., MOLECULAR CLONING: A LABORATORY MANUAL* (2d ed. 1989)). In general, cDNA sequences encoding infectious

Girdwood S.A. or TR339 virus RNA transcripts that hybridize to the cDNAs disclosed herein will be at least 30% homologous, 50% homologous, 75% homologous, and even 95% homologous or more with the cDNA sequences encoding infectious Girdwood S.A. or TR339 virus RNA transcripts disclosed herein.

Promoter sequences and Girdwood S.A. virus or Sindbis virus strain TR339 cDNA clones are operatively associated in the present invention such that the promoter causes the cDNA clone to be transcribed in the presence of an RNA polymerase which binds to the promoter. The promoter is positioned on the 5' end (with respect to the virion RNA sequence), of the cDNA clone. An excessive number of nucleotides between the promoter sequence and the cDNA clone will result in the inoperability of the construct. Hence, the number of nucleotides between the promoter sequence and the cDNA clone is preferably not more than eight, more preferably not more than five, still more preferably not more than three, and most preferably not more than one.

Examples of promoters which are useful in the cDNA sequences of the present invention include, but are not limited to T3 promoters, T7 promoters, cytomegalovirus (CMV) promoters, and SP6 promoters. The DNA sequence of the present invention may reside in any suitable transcription vector. The DNA sequence preferably has a complementary DNA sequence bound thereto so that the double-stranded sequence will serve as an active template for RNA polymerase. The transcription vector preferably comprises a plasmid. When the DNA sequence comprises a plasmid, it is preferred that a unique restriction site be provided 3' (with respect to the virion RNA sequence) to the cDNA clone. This provides a means for linearizing the DNA sequence to allow the transcription of genome-length RNA *in vitro*.

The cDNA clones can be generated by any of a variety of suitable methods known to those skilled in the art. A preferred method is the method set forth in United States Patent No. 5,185,440 to Davis et al., the disclosure of which

is incorporated in its entirety by reference, and Gubler et al., *Gene* 25:263 (1983).

RNA is preferably synthesized from the DNA sequence *in vitro* using purified RNA polymerase in the presence of ribonucleotide triphosphates and cap analogs in accordance with conventional techniques. However, the RNA may also be synthesized intracellularly after introduction of the cDNA.

The Girdwood S.A. and TR339 cDNA clones and the infectious RNAs and infectious virus particles produced therefrom of the present invention are useful for the preparation of pharmaceutical formulations, such as vaccines. In addition, the cDNA clones, infectious RNAs, and infectious viral particles of the present invention are useful for administration to animals for the purpose of producing antibodies to the Girdwood S.A. virus or the Sindbis virus strain TR339, which antibodies may be collected and used in known diagnostic techniques for the detection of Girdwood S.A. virus or Sindbis virus strain TR339. Antibodies can also be generated to the viral proteins expressed from the cDNAs disclosed herein. As another aspect of the present invention, the claimed cDNA clones are useful as nucleotide probes to detect the presence of Girdwood S.A. or TR339 genomic RNA or transcripts.

III. Infectious Virus Particles and Pharmaceutical Formulations.

The infectious virus particles of the present invention include those containing double promoter vectors and those containing replicon vectors as described hereinabove. Alternately, the infectious virus particles contain infectious RNAs encoding the Girdwood S.A. or TR339 genome. When the infectious RNA comprises the Girdwood S.A. genome, preferably the RNA has the sequence encoded by the cDNA given as SEQ ID NO:4. When the infectious RNA comprises the TR339 genome, preferably the RNA has the sequence encoded by the cDNA given as SEQ ID NO:8.

The infectious, alphavirus particles of the present invention may be prepared according to the methods disclosed herein in combination with techniques

-29-

known to those skilled in the art. These methods include transfecting an alphavirus-permissive cell with a replicon RNA including the alphavirus packaging segment and an inserted heterologous RNA, a first helper RNA including RNA encoding at least one alphavirus structural protein, and a second helper RNA including RNA encoding at least one alphavirus structural protein which is different from that encoded by the first helper RNA. Alternately, and preferably, at least one of the helper RNAs is produced from a cDNA encoding the helper RNA and operably associated with an appropriate promoter, the cDNA being stably transfected and integrated into the cells. More preferably, all of the helper RNAs will be "launched" from stably transfected cDNAs. The step of transfecting the alphavirus-permissive cell can be carried out according to any suitable means known to those skilled in the art, as described above with respect to propagation-competent viruses.

Uptake of propagation-competent RNA into the cells *in vitro* can be carried out according to any suitable means known to those skilled in the art. Uptake of RNA into the cells can be achieved, for example, by treating the cells with DEAE-dextran, treating the RNA with LIPOFECTIN® before addition to the cells, or by electroporation, with electroporation being the currently preferred means. These techniques are well known in the art. *See e.g.*, United States Patent No. 5,185,440 to Davis et al., and PCT Publication No. WO 92/10578 to Bioption AB, the disclosures of which are incorporated herein by reference in their entirety. Uptake of propagation-competent RNA into the cell *in vivo* can be carried out by administering the infectious RNA to a subject as described in Section I above.

The infectious RNAs may also contain a heterologous RNA segment, where the heterologous RNA segment contains a heterologous RNA and a promoter operably associated therewith. It is preferred that the infectious RNA introduces and expresses the heterologous RNA in bone marrow cells as described in Section I above. According to this embodiment, it is preferable that the promoter operatively associated with the heterologous RNA is operable in bone

marrow cells. The heterologous RNA may encode any protein or peptide, preferably an immunogenic protein or peptide, a therapeutic protein or peptide, a hormone, a growth factor, an interleukin, a cytokine, a chemokine, an enzyme, a ribozyme, or an antisense oligonucleotide as described in more detail in Section I above.

The step of facilitating the production of the infectious viral particles in the cells may be carried out using conventional techniques. *See e.g.*, United States Patent No. 5,185,440 to Davis et al., PCT Publication No. WO 92/10578 to Bioption AB, and United States Patent No. 4,650,764 to Temin et al. (although Temin et al., relates to retroviruses rather than alphaviruses). The infectious viral particles may be produced by standard cell culture growth techniques.

The step of collecting the infectious virus particles may also be carried out using conventional techniques. For example, the infectious particles may be collected by cell lysis, or collection of the supernatant of the cell culture, as is known in the art. *See e.g.*, United States Patent No. 5,185,440 to Davis et al., PCT Publication No. WO 92/10578 to Bioption AB, and United States Patent No. 4,650,764 to Temin et al. Other suitable techniques will be known to those skilled in the art. Optionally, the collected infectious virus particles may be purified if desired. Suitable purification techniques are well known to those skilled in the art.

Pharmaceutical formulations, such as vaccines, of the present invention comprise an immunogenic amount of the infectious, virus particles in combination with a pharmaceutically acceptable carrier. An "immunogenic amount" is an amount of the infectious virus particles which is sufficient to evoke an immune response in the subject to which the pharmaceutical formulation is administered. An amount of from about 10^3 to about 10^7 particles, and preferably about 10^4 to 10^6 particles per dose is believed suitable, depending upon the age and species of the subject being treated, and the immunogen against which the immune response is desired.

Pharmaceutical formulations of the present invention for therapeutic use comprise a therapeutic amount of the infectious virus particles in combination with a pharmaceutically acceptable carrier. A "therapeutic amount" is an amount of the infectious virus particles which is sufficient to produce a therapeutic effect (e.g., triggering an immune response or supplying a protein to a subject in need thereof) in the subject to which the pharmaceutical formulation is administered. The therapeutic amount will depend upon the age and species of the subject being treated, and the therapeutic protein or peptide being administered. Typical dosages are an amount from about 10^1 to about 10^5 infectious units.

Exemplary pharmaceutically acceptable carriers include, but are not limited to, sterile pyrogen-free water and sterile pyrogen-free physiological saline solution. Subjects which may be administered immunogenic amounts of the infectious virus particles of the present invention include but are not limited to human and animal (e.g., pig, cattle, dog, horse, donkey, mouse, hamster, monkeys) subjects.

Pharmaceutical formulations of the present invention include those suitable for parenteral (e.g., subcutaneous, intracerebral, intradermal, intramuscular, intravenous and intraarticular) administration. Alternatively, pharmaceutical formulations of the present invention may be suitable for administration to the mucus membranes of a subject (e.g., intranasal administration by use of a dropper, swab, or inhaler). The formulations may be conveniently prepared in unit dosage form and may be prepared by any of the methods well known in the art.

The following examples are provided to illustrate the present invention, and should not be construed as limiting thereof. In these examples, PBS means phosphate buffered saline, EDTA means ethylene diamine tetraacetate, ml means milliliter, μ l means microliter, mM means millimolar, μ M means micromolar, u means unit, PFU means plaque forming units, g means gram, mg means milligram, μ g means microgram, cpm means counts per minute, ic means

-32-

intracerebral or intracerebrally, ip means intraperitoneal or intraperitoneally, iv means intravenous or intravenously, and sc means subcutaneous or subcutaneously.

Amino acid sequences disclosed herein are presented in the amino to carboxyl direction, from left to right. The amino and carboxyl groups are not presented in the sequence. Nucleotide sequences are presented herein by single strand only in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented herein in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission, or (for amino acids) by either one letter or three letter code, in accordance with 37 CFR § 1.822 and established usage. Where one letter amino acid code is used, the same sequence is also presented elsewhere in three letter code.

EXAMPLE I

Cells and Virus Stocks

S.A.AR86 was isolated in 1954 from a pool of *Culex* sp. mosquitoes collected near Johannesburg, South Africa. Weinbren et al., *S. Afr. Med. J.* 30, 631-36 (1956). Ockelbo82 was isolated from *Culiseta* sp. mosquitoes collected in Edsbyn, Sweden in 1982 and was associated serologically with human disease. Niklasson et al., *Am. J. Trop. Med. Hyg.* 33, 1212-17 (1984). Girdwood S.A. was isolated from a human patient in the Johannesburg area of South Africa in 1963. Malherbe et al., *S. Afr. Med. J.* 37, 547-52 (1963). Molecularly cloned virus TR339 represents the deduced consensus sequence of Sindbis AR339. McKnight et al., *J. Virol.* 70, 1981-89 (1996); William Klimstra, personal communication. TRSB is a laboratory strain of Sindbis isolate AR339 derived from a cDNA clone pTRSB and differing from the AR339 consensus sequence at three codons. McKnight et al., *J. Virol.* 70, 1981-89 (1996). pTR5000 is a full-length cDNA clone of Sindbis AR339 following the SP6 phage promoter and containing mostly Sindbis AR339 sequences.

Stocks of all molecularly cloned viruses were prepared by electroporating genome length *in vitro* transcripts of their respective cDNA clones

-33-

in BHK-21 cells. Heidner et al., *J. Virol.* 68, 2683-92 (1994). Girdwood S.A. (Malherbe et al., *S. Afr. Med. J.* 37, 547-52 (1963)) and Ockelbo82 (Espmark and Niklasson, *Am. J. Trop. Med. Hyg.* 33, 1203-11 (1984); Niklasson et al., *Am. J. Trop. Med. Hyg.* 33, 1212-17 (1984)) were passed one to three times in BHK-21
5 cells in order to produce amplified stocks of virus. All virus stocks were stored at -70°C until needed. The titers of the virus stocks were determined on BHK-21 cells from aliquots of frozen virus.

EXAMPLE 2

Cloning the S.A.AR86 and Girdwood S.A. Genomic Sequences

10 The sequences of S.A.AR86 (Figure 1, SEQ ID NO: 1) and Girdwood S.A. (Figure 3, SEQ ID NO:4) were determined from uncloned reverse transcriptase-polymerase chain reaction (RT-PCR) fragments amplified from virion RNA. Heidner et al., *J. Virol.* 68, 2683-92 (1994). The sequence of the 5' 40
15 nucleotides was determined by directly sequencing the genomic RNA. Sanger et al., *Proc. Natl. Acad. Sci. USA* 74, 5463-67 (1977); Zimmern and Kaesberg, *Proc. Natl. Acad. Sci. USA* 75, 4257-61 (1978); Ahlquist et al., *Cell* 23, 183-89 (1981).

The S.A.AR86 genome was 11,663 nucleotides in length, excluding the 5' CAP and 3' poly(A) tail, 40 nucleotides shorter than the alphavirus prototype
20 Sindbis strain AR339. Strauss et al., *Virology* 133, 92-110 (1984). Compared with the consensus sequence of Sindbis virus AR339 (McKnight et al., *J. Virol.* 70 1981-89 (1996)), S.A.AR86 contained two separate 6-nucleotide insertions, and one 3-nucleotide insertion in the 3' half of the nsP3 gene, a region not well conserved among alphaviruses. The two 6-nucleotide insertions were found
25 immediately 3' of nucleotides 5403 and 5450, and the 3-nucleotide insertion was immediately 3' of nucleotide 5546 compared with the AR339 genome. In addition, S.A.AR86 contained a 54-nucleotide deletion in nsP3 which spanned nucleotides 5256 to 5311 of AR339. As a result of these deletions and insertions, S.A.AR86 nsP3 was 13 amino acids smaller than AR339, containing an 18-amino acid
30 deletion and a total of 5 amino acids inserted. The 3' untranslated region of

-34-

S.A.AR86 contained, with respect to AR339, two 1-nucleotide deletions at nucleotides 11,513 and 11,602, and one 1-nucleotide insertion following nucleotide 11,664. The total numbers of nucleotides and predicted amino acids comprising the remaining genes of S.A.AR86 were identical to those of AR339.

5 A notable feature of the deduced amino acid sequence of S.A.AR86 (Figure 2, SEQ ID NO:2 and SEQ ID NO:3) was the cysteine codon in place of an opal termination codon between nsP3 and nsP4. S.A.AR86 is the only alphavirus of the Sindbis group, and one of just three alphavirus isolates sequenced to date, which do not contain an opal termination codon between nsP3 and nsP4.
10 Takkinen, K., *Nucleic Acids Res.* 14, 5667-5682 (1986); Strauss et al., *Virology* 164, 265-74 (1988).

 The genome of Girdwood S.A. was 11,717 nucleotides long excluding the 5' CAP and 3' poly(A) tail. The nucleotide sequence (SEQ ID NO:4) of the Girdwood S.A. genome and the putative amino acid sequence (SEQ
15 ID NO:5 and SEQ ID NO:6) of the Girdwood S.A. gene products are shown in Figure 3 and Figure 4, respectively. The asterisk at position 1902 in SEQ ID NO:5 indicates the position of the opal termination codon in the coding region of the nonstructural polyprotein. The extra nucleotides relative to AR339 were in the nonconserved half of nsP3, which contained insertions totalling 15 nucleotides, and
20 in the 3' untranslated region which contained two 1-nucleotide deletions and a 1-nucleotide insertion with respect to the consensus Sindbis AR339 genome. The insertions found in the nsP3 gene of Girdwood S.A. were identical in position and content to those found in S.A.AR86, although Girdwood S.A. did not have the large nsP3 deletion characteristic of S.A.AR86. The remaining portions of the
25 genome contained the same number of nucleotides and predicted amino acids as Sindbis AR339.

 Overall, Girdwood S.A. was 94.5% identical to the consensus Sindbis AR339 sequence, differing at 655 nucleotides not including the insertions and deletions. These nucleotide differences resulted in 88 predicted amino acid

-35-

changes or a difference of 2.3%. A plurality of amino acid differences were concentrated in the nsP3 gene, which contained 32 of the amino acid changes, 25 of which were in the nonconserved 3' half.

5 The Girdwood S.A. nucleotides at positions 1, 3, and 11,717 could not be resolved. Because the primer used during the RT-PCR amplification of the 3' end of the genome assumed a cytosine in the 3' terminal position, the identity of this nucleotide could not be determined with certainty. However, in all alphaviruses sequenced to date there is a cytosine in this position. This, combined with the fact that no difficulty was encountered in obtaining RT-PCR product for
10 this region with an oligo(dT) primer ending with a 3'G, suggested that Girdwood S.A. also contains a cytosine at this position. The ambiguity at nucleotide positions 1 and 3 resulted from strong stops encountered during the RNA sequencing.

EXAMPLE 3

Comparison of S.A.AR86 and Girdwood S.A.

Sequences With Other Sindbis-Related Virus Sequences

15 Table 1 examines the relationship of S.A.AR86 and Girdwood S.A. to each other and to other Sindbis-related viruses. This was accomplished by aligning the nucleotide and deduced amino acid sequences of Ockelbo82, AR339 and Girdwood S.A. to those of S.A.AR86 and then calculating the percentage
20 identity for each gene using the programs contained within the Wisconsin GCG package (Genetics Computer Group, 575 Science Drive, Madison WI 53711); as described in more detail in McKnight et al., *J. Virol.* 70, 1981-89 (1996).

25 The analysis suggests that S.A.AR86 is most similar to the other South African isolate, Girdwood S.A., and that the South African isolates are more similar to the Swedish Ockelbo82 isolate than to the Egyptian Sindbis AR339 isolate. These results also suggest that it is unlikely that S.A.AR86 is a recombinant virus like WEE virus. Hahn et al., *Proc. Natl. Acad. Sci. USA* 85, 5997-6001 (1988).

TABLE 1
Comparison of the Nucleotide and Amino Acid Sequences
of S.A.AR86 Virus with Those of Sindbis AR339, Ockelbo82, and Girdwood S.A. Viruses^a

Regions	Nucleotide Differences ^b			Amino Acid Differences ^b		
	AR339	OCK82	GIRD	AR339	OCK82	GIRD
	Number (%)			Number (%)		
5' untranslated	0 (0.0)	0 (0.0)	1 (1.7)	--	--	--
nsP1	76 (4.7)	37 (2.3)	15 (0.9)	9 (1.7)	6 (1.1)	2 (0.4)
nsP2	137 (5.7)	86 (3.6)	45 (1.9)	15 (1.9)	8 (1.0)	12 (1.5)
nsP3						
Conserved ^c	51 (5.7)	35 (3.9)	13 (1.6)	6 (2.0)	1 (0.3)	1 (0.4)
Nonconserved ^d	116 (6.6)	83 (4.4)	70 (2.2)	45 (9.7)	34 (7.0)	27 (3.7)
nsP4	111 (6.1)	68 (3.7)	19 (1.1)	8 (1.3)	2 (0.3)	4 (0.6)
26s junction	1 (2.1)	0 (0.0)	1 (2.1)	--	--	--
Capsid	36 (4.5)	26 (3.3)	7 (0.9)	1 (0.4)	3 (1.1)	0 (0.0)
E3	17 (8.9)	5 (2.6)	4 (2.1)	1 (1.6)	0 (0.0)	0 (0.0)
E2	71 (5.6)	43 (3.4)	18 (1.4)	12 (2.6)	6 (1.4)	2 (0.5)
6K	10 (6.1)	9 (5.4)	4 (2.4)	2 (3.6)	2 (3.6)	1 (1.8)
E1	49 (3.7)	31 (2.3)	16 (1.2)	7 (1.6)	6 (1.4)	2 (0.9)
3' untranslated	14 (4.5)	8 (2.5)	1 (0.3)	--	--	--
Totals	689 (5.5)	431 (3.3)	214 (1.4)	106 (2.3)	68 (1.4)	51 (0.9)

a. All nucleotide positions and gene boundaries are numbered according to those used for the Sindbis AR339, HR_{sp} variant Genebank Accession No. J02363; Strauss et al., *Virology* 133, 92-110 (1984).

b. Differences include insertions and deletions.

c. Conserved region nucleotides 4100 to 5000 (aa 1 to aa300).

d. Nonconserved region nucleotides 5001 to 5729 (aa301 to aa542, S.A.AR86 numbering).

-37-

EXAMPLE 4

Neurovirulence of S.A.AR86 and Girdwood S.A.

Girdwood S.A., Ockelbo82, and S.A.AR86 are related by sequence; in contrast, it has previously been reported that only S.A.AR86 displayed the adult mouse neurovirulence phenotype. Russell et al., *J. Virol.* 63, 1619-29 (1989). These findings were confirmed by the present investigations. Briefly, groups of four female CD-1 mice (3-6 weeks of age) were inoculated ic with 10^3 plaque-forming units (PFU) of S.A.AR86, Girdwood S.A., or Ockelbo82. Neither Girdwood S.A. nor Ockelbo82 infection produced any clinical signs of infection. Infection with S.A.AR86 produced neurological signs within four to five days and ultimately killed 100% of the mice as previously demonstrated.

Table 2 lists those amino acids of S.A.AR86 which might explain the neurovirulence phenotype in adult mice. A position was scored as potentially related to the S.A.AR86 adult neurovirulence phenotype if the S.A.AR86 amino acid differed from that which otherwise was absolutely conserved at that position in the other viruses.

TABLE 2

Divergent Amino Acids in S.A.AR86
Potentially Related to the Adult Neurovirulence Phenotype

	Position in S.A.AR86	S.A.AR86 Amino Acid	Conserved Amino Acid
nsP1	583	Thr	Ile
nsP2	256	Arg	Ala
	648	Ile	Val
	651	Lys	Glu
nsP3	344	Gly	Glu
	386	Tyr	Ser
	441	Asp	Gly
	445	Ile	Met
	537	Cys	Opal
E2	243	Ser	Leu
6K	30	Val	Ile
E1	112	Val	Ala
	169	Leu	Ser

-38-

EXAMPLE 5

pS55 Molecular Clone of S.A.AR86

As a first step in investigating the unique adult mouse neurovirulence phenotype of S.A.AR86, a full-length cDNA clone of the S.A.AR86 genome was constructed. The sources of cDNA included conventional cDNA clones (Davis et al., *Virology* 171, 189-204 (1989)) as well as uncloned RT-PCR fragments derived from the S.A.AR86 genome. As described previously, these were substituted, starting at the 3' end, into pTR5000 (McKnight et al., *J. Virol.* 70, 1981-89 (1996)), a full-length Sindbis clone from which infectious genomic replicas could be derived by transcription with SP6 polymerase *in vitro*.

The end result was pS55, a molecular clone of S.A.AR86 from which infectious transcripts could be produced and which contained four nucleotide changes (G for A at nt 215; G for C at nt 3863; G for A at nt 5984; and C for T at nt 9113) but no amino acid coding differences with respect to the S.A.AR86 genomic RNA (amino acid sequence of S.A.AR86 presented in Figure 2 (SEQ ID NO:2 and SEQ ID NO:3)). The nucleotide sequence of clone pS55 is presented in Figure 5 (SEQ ID NO:7).

As has been described by Simpson et al., *Virology* 222, 464-69 (1996), neurovirulence and replication of the virus derived from pS55 (S55) were compared with those of S.A.AR86. It was found that S55 exhibits the distinctive adult neurovirulence characteristic of S.A.AR86. Like S.A.AR86, S55 produces 100% mortality in adult mice infected with the virus and the survival times of animals infected with both viruses were indistinguishable. In addition, S55 and S.A.AR86 were found to replicate to essentially equivalent titers *in vivo*, and the profiles of S55 and S.A.AR86 virus growth in the central nervous system and periphery were very similar.

From these data it was concluded that the silent changes found in virus derived from clone pS55 had little or no effect on its growth or virulence, and that this molecularly cloned virus accurately represents the biological isolate, S.A.AR86.

EXAMPLE 6

Construction of the Consensus AR339 Virus TR339

The consensus sequence of the Sindbis virus AR339 isolate, the prototype alphavirus was deduced. The consensus AR339 sequence was inferred by comparison of the TRSB sequence (a laboratory-derived AR339 strain) with the complete or partial sequences of HR_{sp} (the Gen Bank sequence; Strauss et al., *Virology* 133, 92-110 (1984)), SV1A, and NSV (AR339-derived laboratory strains; Lustig et al., *J. Virol* 62, 2329-36 (1988)), and SIN (a laboratory-derived AR339 strain; Davis et al., *Virology* 161, 101-108 (1987), Strauss et al., *J. Virol.* 65, 4654-64 (1991)). Each of these viruses was descended from AR339. Where these sequences differed from each other, they also were compared with the amino acid sequences of other viruses related to Sindbis virus: Ockelbo82, S.A.AR86, Girdwood S.A., and the somewhat more distantly related Aura virus. Rumenapf et al., *Virology* 208, 621-33 (1995).

The details of determining a consensus AR339 sequence and constructing the consensus virus TR339 have been described elsewhere. McKnight et al., *J. Virol.* 70, 1981-89 (1996); Klimstra et al., *manuscript in preparation*. The nucleotide (SEQ ID NO:8) sequence of pTR339 is presented in Figure 6. The deduced amino acid sequences of the pTR339 non-structural and structural polyproteins are shown as SEQ ID NO:9 and SEQ ID NO:10, respectively. The asterisk at position 1897 in SEQ ID NO:9 indicates the position of the opal termination codon in the coding region of the nonstructural polyprotein. The consensus nucleotide sequence diverged from the pTRSB sequence at three coding positions (nsP3 528, E2 1, and E1 72). These differences are illustrated in Table 3.

TABLE 3

Amino Acid Differences Between
Laboratory Strain TRSB and Molecular Clone TR339

	nsP3 528 (nt5683)	E2 1 (nt8633)	E1 72 (nt10279)
TR339	Arg (CGA)	Ser (AGC)	Ala (GCU)
TRSB	Gln (CAA)	Arg (AGA)	Val (GUU)

EXAMPLE 7

Animals Used for *In Vivo* Localization Studies

Specific pathogen free CD-1 mice were obtained from Charles River Breeding Laboratories (Raleigh, North Carolina) at 21 days of age and maintained under barrier conditions until approximately 37 days of age. Intracerebral (ic) inoculations were performed as previously described, Simpson et al., *Virol.* 222, 464-49 (1996), with 500 PFU of S51 (an attenuated mutant of S55) or 10^3 PFU of S55. Animals inoculated peripherally were first anesthetized with METOFANE®. Then, 25 μ l of diluent (PBS, pH 7.2, 1% donor calf serum, 100 u/ml penicillin, 50 μ g/ml streptomycin, 0.9 mM CaCl_2 , and 0.5 mM MgCl_2) containing 10^3 PFU of virus were injected either intravenously (iv) into the tail vein, subcutaneously (sc) into the skin above the shoulder blades on the middle of the back, or intraperitoneally (ip) in the lower right abdomen. Animals were sacrificed at various times post-inoculation as previously described. Simpson et al., *Virol.* 222, 464-49 (1996). Brains (including brainstems) were homogenized in diluent to 30% w/v, and right quadriceps were homogenized in diluent to 25% w/v. Homogenates were handled and titered as described previously. Simpson et al., *Virol.* 222, 464-49 (1996). Bone marrow was harvested by crushing both femurs from each animal in sufficient diluent to produce a 30% w/v suspension (calculated as weight of uncrushed femurs in volume of diluent). Samples were stored at -70°C . For titration, samples were thawed and clarified by centrifugation at $1,000 \times g$ for 20 minutes at 4°C before being titered by conventional plaque assay on BHK-21 cells.

EXAMPLE 8

Tissue Preparation for *In Situ* Hybridization Studies

Animals were anesthetized by ip injection of 0.5 ml AVERTIN® at various times post-inoculation followed by perfusion with 60 to 75 ml of 4% paraformaldehyde in PBS (pH 7.2) at a flow rate of 10 ml per minute. The entire carcass was decalcified for 8 to 10 weeks in 4% paraformaldehyde containing 8% EDTA in PBS (pH 6.8) at 4°C . This solution was changed twice during the decalcification period. Selected tissues were cut into blocks approximately 3 mm thick and placed into biopsy cassettes for paraffin embedding and sectioning. Blocks were embedded, sectioned and hematoxylin/eosin stained by Experimental Pathology Laboratories (Research Triangle Park, North Carolina) or North

Carolina State University Veterinary School Pathology Laboratory (Raleigh, North Carolina).

EXAMPLE 9

In Situ Hybridization

Hybridizations were performed using a [³⁵S]-UTP labeled S.A.AR86 specific riboprobe derived from pDS-45. Clone pDS-45 was constructed by first amplifying a 707 base pair fragment from pS55 by PCR using primers 7241 (5'-CTGCGGCGGATTCATCTTGC-3', SEQ ID NO:11) and SC-3 (5'-CTCCAACTTAAGTG-3', SEQ ID NO:12). The resulting 707 base pair fragment was purified using a GENE CLEAN® kit (Bio101, CA), digested with *Hha*I, and cloned into the *Sma*I site of pSP72 (Promega). Linearizing pDS-45 with *Eco*RV and performing an *in vitro* transcription reaction with SP6 DNA-dependent, RNA polymerase (Promega) in the presence of [³⁵S]-UTP resulted in a riboprobe approximately 500 nucleotides in length of which 445 nucleotides were complementary to the S.A.AR86 genome (nucleotides 7371 through 7816). A riboprobe specific for the influenza strain PR-8 hemagglutinin (HA) gene was used as a control probe to test non-specific binding. The *in situ* hybridizations were performed as described previously (Charles et al., *Virology* 208, 662-71 (1995)) using 10⁵ cpm of probe per slide.

EXAMPLE 10

Replication of S.A.AR86 in Bone Marrow

Three groups of six adult mice each were inoculated peripherally (sc, ip, or iv) with 1200 PFU of S55 (a molecular clone of S.A.AR86) in 25 µl of diluent. Under these conditions, the infection produced no morbidity or mortality. Two mice from each group were anesthetized and sacrificed at 2, 4 and 6 days post-inoculation by exsanguination. The serum, brain (including brainstem), right quadricep, and both femurs were harvested and titered by plaque assay. Virus was never detected in the quadricep samples of animals inoculated sc (Table 4). A single animal inoculated ip (two days post-inoculation) and two mice inoculated iv (at four and six days post-inoculation) had detectable virus in the right quadricep, but the titer was at or just above the limit of detection (6.25 PFU/g tissue). Virus was present sporadically or at low levels in the brain and

serum of animals regardless of the route of inoculation. Virus was detected in the bone marrow of animals regardless of the route of inoculation. However, the presence of virus in bone marrow of animals inoculated sc or ip was more sporadic than animals inoculated iv, where five out of six animals had detectable virus. These results suggest that S55 targets to the bone marrow, especially following iv inoculation.

The level and frequency of virus detected in the serum and muscle suggested that virus detected in the bone marrow was not residual virus contamination from blood or connective tissue remaining in bone marrow samples. The following experiment also suggested that virus in bone marrow was not due to tissue or serum contamination. Mice were inoculated ic with 1200 PFU of S55 in 25 μ l of diluent. Animals were sacrificed at 0.25, 0.5, 1, 1.5, 2, 3, 4, 5, and 6 days post-inoculation, and the carcasses were decalcified as described in Example 8. Coronal sections taken at approximately 3 mm intervals through the head, spine (including shoulder area), and hips were probed with an S55-specific [³⁵S]-UTP labeled riboprobe derived from pDS-45. Positive *in situ* hybridization signal was detected by one day post-inoculation in the bone marrow of the skull (data not shown). Weak signal also was present in some of the chondrocytes of the vertebrae, suggesting that S55 was replicating in these cells as well. Although the frequency of positive bone marrow cells was low, the signal was very intense over individual positive cells. This result strongly suggests that S55 replicates *in vivo* in a subset of cells contained in the bone marrow.

EXAMPLE 11

Other Sindbis Group Viruses

It was of interest to determine if the ability to replicate in the bone marrow of mice was unique to S55 or was a general feature of other viruses, both Sindbis and non-Sindbis viruses, in the Sindbis group. Six 38-day-old female CD-1 mice were inoculated iv with 25 μ l of diluent containing 10³ PFU of S55, Ockelbo82, Girdwood S.A., TR339, or TRSB. At 2, 4 and 6 days post-inoculation two mice from each group were sacrificed and whole blood, serum, brain (including brainstem), right quadricep, and both femurs were harvested for virus titration.

The results of this experiment were similar to those with S55. TRSB infected animals had no virus detectable in serum or whole blood in any animal at any time, and with the other viruses tested, no virus was detected in the serum or whole blood of any animal beyond two days post-inoculation (detection limit, 25 PFU/ml). Neither TRSB nor TR339 was detectable in the brains of infected animals at any time post-inoculation. S55, Girdwood S.A., and Ockelbo82 were present in the brains of infected animals sporadically with the titers being at or near the 75 PFU/g level of detection. All the tested viruses were found sporadically at or slightly above the 50 PFU/g detection limit in the right quadricep of infected animals except for a single animal four days post-inoculation with TRSB which had nearly 10^5 PFU/g of virus in its quadricep.

The frequency at which the different viruses were detected in bone marrow varied widely, with S55 and Girdwood S.A. being the most frequently isolated (five out of six animals) and Ockelbo82 and TRSB being the least frequently isolated from bone marrow (one out of six animals and two out of six animals, respectively) (Table 4). Girdwood S.A. and S55 gave nearly identical profiles in all tissues. Girdwood S.A., unlike S.A.AR86, is not neurovirulent in adult mice (Example 4), suggesting that the adult neurovirulence phenotype is distinct from the ability of the virus to replicate efficiently in bone marrow.

TABLE 4
Titers Following IV Inoculation of Virus

Tissue Titered								
Virus	Animal	Days Post-Inoculation	Bone Marrow (PFU/g)	Serum (PFU/ml)	Blood (PFU/ml)	Brain (PFU/g)	Quadricep (PFU/g)	
S55	A	2	1125	N.D.*	N.D.	N.D.	N.D.	
	B		488	50	200	N.D.	N.D.	
	A	4	863	N.D.	N.D.	N.D.	550	
	B		113	N.D.	N.D.	75	N.D.	
	A	6	N.D.	N.D.	N.D.	N.D.	50	
	B		37.5	N.D.	N.D.	N.D.	N.D.	
	Limit of Detection			37.5	25	75	50	
	TR339	A	2	N.D.	N.D.	N.D.	N.D.	N.D.
		B		1500	75	700	N.D.	N.D.
		A	4	1050	N.D.	N.D.	N.D.	N.D.
B		1762		N.D.	N.D.	N.D.	400	
A		6	N.D.	N.D.	N.D.	N.D.	N.D.	
B			N.D.	N.D.	N.D.	N.D.	N.D.	
Limit of Detection			37.5	25	37.5	50		
TRSB		A	2	N.D.	N.D.	N.D.	N.D.	N.D.
		B		N.D.	N.D.	N.D.	N.D.	N.D.
		A	4	150	N.D.	N.D.	N.D.	1000
	B	N.D.		N.D.	N.D.	N.D.	100000	
	A	6	N.D.	N.D.	N.D.	N.D.	N.D.	
	B		37.5	N.D.	N.D.	N.D.	N.D.	
	Limit of Detection			37.5	25	37.5	50	

TABLE 4 Continued
Titers Following IV Inoculation of Virus

Tissue Titered							
Virus	Animal	Days Post-Inoculation	Bone Marrow (PFU/g)	Serum (PFU/ml)	Blood (PFU/ml)	Brain (PFU/g)	Quadricep (PFU/g)
Girdwood S.A.	A	2	22000	2325	1450	30 0	50
	B		2500	1200	2600	N.D.	N.D.
	A	4	788	N.D.	N.D.	N.D.	N.D.
	B		113	N.D.	N.D.	75	N.D.
	A	6	N.D.	N.D.	N.D.	N.D.	N.D.
	B		75	N.D.	N.D.	1700	N.D.
	Limit of Detection		37.5	25	25	75	50
	A	2	N.D.	125	150	N.D.	N.D.
	B		N.D.	50	500	N.D.	200
	A	4	N.D.	N.D.	N.D.	300	N.D.
Ockelbo82	B		300	N.D.	N.D.	N.D.	N.D.
	A	6	N.D.	N.D.	N.D.	100000	N.D.
	B		N.D.	N.D.	N.D.	N.D.	N.D.
	Limit of Detection		37.5	25	25	75	50

* "N.D." indicates that the virus titers were below the limit of detection.

EXAMPLE 12

Virus Persistence in Bone Marrow

The next step in our investigations was to evaluate the possibility that S.A.AR86 persisted long-term in bone marrow. S51 is a molecularly cloned, attenuated mutant of S55. S51 differs from S55 by a threonine for isoleucine substitution at amino acid residue 538 of nsP1 and is attenuated in adult mice inoculated intracerebrally. Like S55, S51 targeted to and replicated in the bone marrow of 37-day-old female CD-1 mice following ic inoculation. Mice were inoculated ic with 500 PFU of S51 and sacrificed at 4, 8, 16, and 30 days post-inoculation for determination of bone marrow and serum titers. At no time post-inoculation was virus detected in the serum above the 6.25 PFU/ml detection limit. Virus was detectable in the bone marrow samples of both animals sampled at four days post-inoculation and in one animal eight days post-inoculation (Table 5). No virus was detectable by titration on BHK-21 cells in any of the bone marrow samples beyond eight days post-inoculation. These results suggested that the attenuating mutation present in S51, which reduces the neurovirulence of the virus, did not impair acute viral replication in the bone marrow.

It was notable that the plaque size on BHK-21 cells of virus recovered on day 4 post-inoculation was smaller than the size of plaques produced by the inoculum virus, and that plaques produced from virus recovered from the day 8 post-inoculation samples were even smaller and barely visible. This suggests a strong selective pressure in the bone marrow for virus that is much less efficient in forming plaques on BHK-21 cells.

To demonstrate that S51 virus genomes were present in bone marrow cells long after acute infection, four to six-week-old female CD-1 mice were inoculated ic with 500 PFU of S51. Three months post-inoculation two animals were sacrificed, perfused with paraformaldehyde and decalcified as described in Example 8. The heads and hind limbs from these animals were paraffin embedded, sectioned, and probed with a S.A.AR86 specific [³⁵S]-UTP labeled riboprobe derived from clone pDS-45. *In situ* hybridization signal was clearly present in discrete cells of the bone and bone marrow of the legs (data not shown). Furthermore, no *in situ* hybridization signal was detected in an adjacent

-47-

control section probed with an influenza virus HA gene specific riboprobe. As the relative sensitivity of *in situ* hybridization is reduced in decalcified tissues (Peter Charles, personal communication), these cells likely contain a relatively high number of viral sequences, even at three months post-inoculation. No *in situ* hybridization signal was observed in mid-sagittal sections of the heads with the S.A.AR86 specific probe, although focal lesions were observed in the brain indicative of the prior acute infection with S51.

TABLE 5

S51 Titers in Bone Marrow Following IC Inoculation of 500 PFU			
Days Post-Inoculation	Titers (Total PFU/Animal)		Limit of Detection
	Animal A	Animal B	
4	2100	380	62.5
8	62.5	N.D. ^a	62.5
16	N.D.	N.D.	62.5
30	N.D.	N.D.	62.5

^a "N.D." indicates that the virus titers were below the limit of detection.

Example 13

Replication of S.A.A.R86 within Bone/Joint Tissue of Adult Mice

Several old world alphaviruses, including Ross River Virus, Chikungunya virus, Okelbo82, and S.A.AR86 are associated with acute and persistent
5 arthritis/arthralgia in humans. Molecular clones of several Sindbis group viruses, including S.A.AR86, were used to investigate alphavirus replication within bone/joint tissue.

Following intravenous inoculation of S.A.AR86 into adult CD-1 mice, viral replication was observed in bone/joint tissue, but not surrounding muscle tissue of
10 the hind limbs. Infectious virus was detectable 24 hrs post-infection; however, viral titer within bone/joint tissue was maximal 72 hours post-infection. Fractionation of hind limbs from infected animals revealed that the hip and knee joints were the predominant sites of viral replication. Replication within bone/joint tissue appears to be a common trait of Sindbis-group viruses, since the laboratory strains TR339 and TRSB
15 also replicated within bone/joint tissue. *In situ* hybridization and S.A.AR86 based double promoter vectors expressing green fluorescent protein were used to further localize S.A.AR86 infected cells within bone/joint tissue. Green fluorescent protein expression was detected in bone/joint tissue for at least one month post-inoculation. These studies demonstrated that cells within the endosteum of synovial joints were the
20 predominant site of S.AAR86 replication.

-49-

SEQUENCE LISTINGS

-50-

THAT WHICH IS CLAIMED IS:

1. A method of introducing and expressing heterologous RNA in bone marrow cells, comprising:

(a) providing a recombinant alphavirus, said alphavirus containing a heterologous RNA segment, said heterologous RNA segment comprising a promoter operable in said bone marrow cells operatively associated with a heterologous RNA to be expressed in said bone marrow cells; and then

(b) contacting said recombinant alphavirus to said bone marrow cells so that said heterologous RNA segment is introduced and expressed therein.

2. A method according to claim 1, wherein said contacting step is carried out *in vitro*.

3. A method according to claim 1, wherein said contacting step is carried out *in vivo* in a subject in need of such treatment.

4. A method according to claim 1, wherein said heterologous RNA encodes a protein or peptide.

5. A method according to claim 1, wherein said heterologous RNA encodes an immunogenic protein or peptide.

6. A method according to claim 1, wherein said heterologous RNA encodes an antisense oligonucleotide or a ribozyme.

7. A method according to claim 1, wherein said alphavirus is an Old World alphavirus.

8. A method according to claim 1, wherein said alphavirus is selected from the group consisting of SF group and SIN group alphaviruses.

9. A method according to claim 1, wherein said alphavirus is selected from the group consisting of Semliki Forest virus, Middelburg virus, Chikungunya virus, O'Nyong-Nyong virus, Ross River virus, Barmah Forest virus, Getah virus, Sagiyama virus, Bebaru virus, Mayaro virus, Una virus, 5 Sindbis virus, South African Arbovirus No. 86, Ockelbo virus, Girdwood S.A. virus, Aura virus, Whataroa virus, Babanki virus, and Kyzylagach virus.

10. A method according to claim 1, wherein said alphavirus is South African Arbovirus No. 86.

11. A method according to claim 1, wherein said alphavirus is 10 Girdwood S.A.

12. A method according to claim 1, wherein said alphavirus is Sindbis strain TR339.

13. A helper cell for expressing an infectious, propagation defective, Girdwood S.A. virus particle, comprising, in a Girdwood S.A.- 15 permissive cell:

(a) a first helper RNA encoding (i) at least one Girdwood S.A. structural protein, and (ii) not encoding at least one other Girdwood S.A. structural protein; and

(b) a second helper RNA separate from said first helper RNA, 20 said second helper RNA (i) not encoding said at least one Girdwood S.A. structural protein encoded by said first helper RNA, and (ii) encoding said at least one other Girdwood S.A. structural protein not encoded by said first helper RNA, and with all of said Girdwood S.A. structural proteins encoded by said first and second helper RNAs assembling together into Girdwood S.A. particles in said cell 25 containing said replicon RNA;

and wherein the Girdwood S.A. packaging segment is deleted from at least said first helper RNA.

-52-

14. The helper cell according to claim 13, further containing a replicon RNA;

said replicon RNA encoding said Girdwood S.A. packaging segment and an inserted heterologous RNA;

5 wherein said Girdwood S.A. packaging segment is deleted from at least one of said helper RNA;

and wherein said replicon RNA, said first helper RNA, and said second helper RNA are all separate molecules from one another.

15. The helper cell according to claim 13, further containing a replicon RNA;

said replicon RNA encoding said Girdwood S.A. packaging segment and an inserted heterologous RNA;

wherein said replicon RNA and said first helper RNA are separate molecules;

15 and wherein the molecule containing said replicon RNA further contains RNA encoding said at least one Girdwood S.A. structural protein not encoded by said first helper RNA.

16. The helper cell according to claim 13, wherein said first helper RNA encodes both the Girdwood S.A. E1 glycoprotein and the Girdwood S.A. E2 glycoprotein, and wherein said second helper RNA encodes the Girdwood S.A. capsid protein.

17. A method of making infectious, propagation defective, Girdwood S.A. virus particles, comprising:

transfecting a Girdwood S.A.-permissive cell according to claim 13 with a propagation defective replicon RNA, said replicon RNA including said Girdwood S.A. packaging segment and an inserted heterologous RNA;

25 producing said Girdwood S.A. virus particles in said transfected cell; and then

collecting said Girdwood S.A. virus particles from said cell.

-53-

18. Infectious Girdwood S.A. virus particles produced by the method of Claim 17.

19. Infectious Girdwood S.A. virus particles containing a replicon RNA encoding a promoter, an inserted heterologous RNA, and wherein RNA encoding at least one Girdwood S.A. structural protein is deleted therefrom so that said Girdwood S.A. virus particle is propagation defective.

20. A pharmaceutical formulation comprising infectious Girdwood S.A. virus particles according to claim 18 or 19 in a pharmaceutically acceptable carrier.

21. A helper cell for expressing an infectious, propagation defective, TR339 virus particle, comprising, in a TR339-permissive cell:

(a) a first helper RNA encoding (i) at least one TR339 structural protein, and (ii) not encoding at least one other TR339 structural protein; and

(b) a second helper RNA separate from said first helper RNA, said second helper RNA (i) not encoding said at least one TR339 structural protein encoded by said first helper RNA, and (ii) encoding said at least one other TR339 structural protein not encoded by said first helper RNA, and with all of said TR339 structural proteins encoded by said first and second helper RNAs assembling together into TR339 particles in said cell containing said replicon RNA;

and wherein the TR339 packaging segment is deleted from at least said first helper RNA.

22. The helper cell according to claim 21, further containing a replicon RNA;

said replicon RNA encoding said TR339 packaging segment and an inserted heterologous RNA;

wherein said TR339 packaging segment is deleted from at least one of said helper RNA;

and wherein said replicon RNA, said first helper RNA, and said second helper RNA are all separate molecules from one another.

-54-

23. The helper cell according to claim 21, further containing a replicon RNA;

said replicon RNA encoding said TR339 packaging segment and an inserted heterologous RNA;

5 wherein said replicon RNA and said first helper RNA are separate molecules;

and wherein the molecule containing said replicon RNA further contains RNA encoding said at least one TR339 structural protein not encoded by said first helper RNA.

10 24. The helper cell according to claim 21, wherein said first helper RNA encodes both the TR339 E1 glycoprotein and the TR339 E2 glycoprotein, and wherein said second helper RNA encodes the TR339 capsid protein.

15 25. A method of making infectious, propagation defective, TR339 virus particles, comprising:

transfecting a TR339-permissive cell according to claim 21 with a propagation defective replicon RNA, said replicon RNA including said TR339 packaging segment and an inserted heterologous RNA;

20 producing said TR339 virus particles in said transfected cell; and then

collecting said TR339 virus particles from said cell.

26. Infectious TR339 virus particles produced by the method of Claim 25.

25 27. Infectious TR339 virus particles containing a replicon RNA encoding a promoter, an inserted heterologous RNA, and wherein RNA encoding at least one TR339 structural protein is deleted therefrom so that said virus particle is propagation defective.

28. A pharmaceutical formulation comprising infectious TR339 virus particles according to Claim 26 or 27 in a pharmaceutically acceptable carrier.

-55-

29. A recombinant DNA comprising a cDNA coding for an infectious Girdwood S.A. virus RNA transcript and a heterologous promoter positioned upstream from said cDNA and operatively associated therewith.

5 30. An infectious RNA transcript encoded by a cDNA according to claim 29.

31. An infectious RNA according to claim 30, said infectious Girdwood S.A. RNA transcript containing a heterologous RNA segment, said heterologous RNA segment comprising a promoter operably associated with a heterologous RNA.

10 32. Infectious viral particles containing an RNA transcript according to claim 30.

33. A recombinant DNA comprising a cDNA coding for a Sindbis strain TR339 RNA transcript and a heterologous promoter positioned upstream from said cDNA and operatively associated therewith.

15 34. An infectious RNA transcript encoded by a cDNA according to claim 33.

20 35. An infectious RNA according to claim 34, said infectious Girdwood S.A. RNA transcript containing a heterologous RNA segment, said heterologous RNA segment comprising a promoter operably associated with a heterologous RNA.

36. Infectious viral particles containing an RNA transcript according to claim 34.

Nucleotide Sequence of S.A.AR86

1 ATTGGCGGCG TAGTACACAC TATTGAATCA AACAGCCGAC CAATTGCACT ACCATCACAA TGGAGAAGCC AGTAGTTAAC GTAGACGTAG ACCCTCAGAG
101 TCCGTTTGTC GTGCAACTGC AAAAGAGCTT CCCGCAATTT GAGGTAGTAG CACAGCAGGT CACTCCAAAT GACCATGCTA ATGCCAGAGC ATTTTCGCAT
201 CTGGCCAGTA AACTAATCGA GCTGGAGGTT CCTACCACAG CGACGATTTT GGACATAGGC AGCGCACCGG CTCGTAGAAT GTTTTCCGAG CACCACTACC
301 ATTGCGTTTG CCCCATGCGT AGTCCAGAAG ACCCGGACCG CATGATGAAA TATGCCAGCA AACTGGCGGA AAAAGCATGT AAGATTACAA ACAAGAATCT
401 GCATGAGAAG ATCAAGGACC TCCGGACCGT ACTTGATACA CCGGATGCTG AAACGCCATC ACTCTGCTTC CACAACGATG TTACCTGCAA CACGCGTGCC
501 GAGTACTCCG TCATGCAGGA CGTGATATC AACGCTCCCG GAACTATTTA CCACCAGGCT ATGAAAGGCG TCGCGACCGT GTACTGGATT GGCTTCGACA
601 CCACCCAGTT CATGTTCTCG GCTATGGCAG GTTCGTACCC TGCATACAA ACCTAACTGG CCGACGAAAA AGTCCTTGAA GCGCGTAACA TCGGACTCTG
701 CAGCACAAAG CTGAGTGAAG GCAGGACAGG AAAGTTGTCT ATAATGAGGA AGAAGGAGTT GAAGCCCGGG TCACGGGTTT ATTTCTCCGT TGGATCGACA
801 CTTTACCCAG AACACAGAGC CAGCTTGAG AGCTGGCAGT TTCCATCGGT GTTCCACTTG AAAGGAAAGC AGTCGTACAC TTGCCGCTGT GATACAGTGG
901 TGAGCTCGA AGGCTACGTA GTGAAGAAAA TCACCATCAG TCCCGGGATC ACGGGAGAAA CCGTGGGATA CGCGGTTACA AACAATAGCG AGGGCTTCTT
1001 GCTATGCAAA GTTACCGATA CAGTAAAGG AGAACGGGTA TCGTTCCCGG TGTGCACGTA TATCCCGGCC ACCATATGCG ATCAGATGAC CGGCATAATG
1101 GCCACGGATA TCTCACCTGA CGATGCACAA AAACCTTCTG TGGGCTCAA CCAGCGAATC GTCATTAACG GTAAGACTAA CAGGAACACC AATACCATGC
1201 AAAATTACCT TCTGCCAATC ATTGCACAAG GGTTCAGCAA ATGGGCCAAG GAGCGCAAG AAGATCTTGA CAATGAAAAA ATGCTGGGCA CCAGAGAGCG
1301 CAAGCTTACA TATGGCTGCT TGTGGCGGTT TCGCACTAAG AAAGTGCACT CGTTCTATCG CCCACCTGGA ACGCAGACCA TCGTAAAAAG CCCAGCCTCT
1401 TTTAGCGCTT TCCCATGTC ATCCGTATGG ACTACCTCTT TGCCCATGTC GCTGAGGCGA AAGATGAAAT TGGCATTACA ACCAAAGAAG GAGGAAAAAC
1501 TGCTGCAAGT CCCGGAGGAA TTAGTTATGG AGGCCAAGGC TGCTTTCGAG GATGCTCAGG AGGAATCCAG AGCGGAGAAG CTCGAGAAG CACTCCACCC
1601 ATTAGTGCCA GACAAAGGTA TCGAGGCAGC TCGGGAAGTT GTCTGCAAG TGGAGGGGCT CCAGGCGGAC ACCGGAGCAG CACTCGTGA AACCCCGCGC
1701 GGTCTGTAA GGATAATACC TCAAGCAAT GACCGTATGA TCGGACAGTA TATCGTTGTC TCGCCGATCT CTGTGCTGAA GAACGCTAAA CTCGCACCAAG
1801 CACACCCGCT AGCAGACCAG GTTAAGATCA TAACGCACTC CGGAAGATCA GGAAGGTATG CAGTCGAACC ATACGACGCT AAAGTACTGA TGCCAGCAGG
1901 AAGTGCGTA CCATGGCCAG AATTCTTAGC ACTGAGTGAG AGCGCCACGC TTGTGTACAA CGAAAGAGAG TTTGTGAACC GCAAGCTGTA CCATATTGCC
2001 ATGCACGGTC CCGCTAAGAA TACAGAAGAG GAGCAGTACA AGGTTACAAA GGCAGAGCTC GCAGAAACAG AGTACGTGTT TGACGTGGAC AAGAAGCGAT
2101 GCGTTAAGAA GGAAGAAGCC TCAGGACTTG TCCTTTCGGG AGAACTGACC AACCCGCCCT ATCAGCAACT AGCTCTTGAG GGACTGAAGA CTCGACCCGC
2201 GGTCCTGAC AAGGTTGAAA CAATAGGAGT GATAGGCACA CCAGGATCGG GCAAGTCAGC TATCATCAAG TCAACTGTCA CGGCACGTGA TCTGTTTACC
2301 AGCGGAAAGA AAGAAAATG CCGCGAAATT GAGGCCGAGC TGCTACGGCT GAGGGGCATG CAGATCACGT CGAAGACAGT GGATTCGGTT ATGCTCAAGC
2401 GATGCCACAA AGCCGTAGAA GTGCTGTATG TTGACGAAGC GTTCCGCTGC CACGCAGGAG CACTACTTGC CTTGATTGCA ATCGTCAGAC CCCGTAAGAA
2501 GGTAGTACTA TCGCGAGACC CTAAGCAATG CGGATTCTTC AACATGATGC AACTAAAGGT ACATTTCAC CACCCTGAAA AAGACATATG TACCAAGACA
2601 TTCTACAAGT TTATCTCCCG ACGTTGCACA CAGCCAGTCA CGGCTATTGT ATCGACACTG CATTACGATG GAAAAATGAA AACCACAAAC CCGTGCAAGA
2701 AGAACATCGA AATCGACATT ACAGGGGCGA CGAAGCCGAA GCCAGGGGAC ATCATCTGTA CATGTTTCCG CGGGTGGGTT AAGCAACTGC AAATCGACTA
2801 TCCCGGACAT GAGGTAATGA CAGCCGCGGC CTCACAAGGG CTAACCAGAA AAGGAGTATA TGCCGTCCCG CAAAAAGTCA ATGAAAACCC GCTGTACGGC
2901 ATCACATCAG AGCATGTGAA CGTGTGCTC ACCCGCACTG AGGACAGGCT AGTATGGAAG ACTTTACAGG GCGACCCATG GATTAAGCAG CTCCTAACG
3001 TACCTAAAGG AAATTTTCAG GCCACCATCG AGGACTGGGA AGCTGAACAC AAGGGAATAA TTGCTGCGAT AAACAGTCCC GCTCCCGTA CCAATCCGTT
3101 CAGCTGCAAG ACTAACGTTT GCTGGCGGAA AGCACTGGAA CCGATACTGG CCACGGCCGG TATCGTACTT ACCGGTTGCC AGTGGAGCGA GCTGTTCCCA
3201 CAGTTTGGCG ATGACAAACC ACCTCGGCC ATCTACGCT TAGACGTAAT TTGCATTAAG TTTTTCGGCA TGGACTTGAC AAGCGGGCTG TTTTCCAAAC
3301 AGAGCATCCC GTTAACGTAC CATCCTGCCG ACTCAGCGAG GCCAGTAGCT CATTGGGACA ACAGCCAGG AACACGCAAG TATGGGTACG ATCAGCCCGT
3401 TGCCGCGGAA CTCTCCCGTA GATTCCCGT GTTCCAGCTA GCTGGGAAAG GCACACAGCT TGATTGCGAG ACGGGCAGAA CTAGAGTTAT CTCTGCACAG
3501 CATAACTTGG TCCAGTGAA CCGCAATCTC CCTCAGCCT TAGTCCCGCA GCACAAGGAG AAACAACCCG GCCCGGTCGA AAAATTCTTG AGCCAGTTCA
3601 AACACCACTC CGTACTGTG ATCTCAGAGA AAAAAATTGA AGTCCCCAC AAGAGAATCG AATGGATCGC CCCGATTGGC ATAGCCGGCG CAGATAAGAA
3701 CTACAACCTG GCTTTCGGGT TTCCGCCGCA GGCACGGTAC GACCTGGTGT TCATCAATAT TGGAACTAAA TACAGAAACC ATCACTTCA ACAGTGCGAA

Fig. 1A

3801 GACCACGCGG CGACCTTGAA AACCTTTTCG CGTTCGGCCC TGAAGTGCCT TAACCCCGGA GGCACCTCG TGGTGAAGTC CTACGGTTAC GCCGACCGCA
3901 ATAGTGAGGA CGTAGTCACC GCTCTTGCCA GAAAAATTGT CAGAGTGTCT GCAGCGAGGC CAGAGTGCCT CTCAGCAAT ACAGAAATGT ACCTGATTTT
4001 CCGACAATA GACAACAGCC GCACACGACA ATCACCCCG CATCATTTGA ATTGTGTGAT TTCGTCCGTG TACGAGGGTA CAAGAGACGG AGTTGGAGCC
4101 GCACCGTCGT ACCGTACTAA AAGGGAGAAC ATTGCTGATT GTCAAGAGGA AGCAGTTGTC AATGCAGCCA ATCCACTGGG CAGACCAGGA GAAGGAGTCT
4201 GCCGTGCCAT CTATAAACGT TGGCCGAACA GTTTCACCGA TTCAGCCACA GAGACAGGTA CCGCAAACT GACTGTGTGC CAAGGAAAGA AAGTGATCCA
4301 CGCGGTTGGC CTGTATTTC GGAACACCC AGAGGCAGAA GCCCTGAAAT TGCTGCAAAA CGCTACCAT GCAGTGGCAG ACTTAGTAAA TGAACATAAT
4401 ATCAAGTCTG TCGCCATCCC ACTGCTATCT ACAGGCATTT ACGCAGCCGG AAAAGACCCG CTTGAGGTAT CACTTAACTG CTTGACAACC GCGCTAGACA
4501 GAACTGATGC GGACGTAACC ATCTACTGCC TGGATAAGAA GTGGAAAGGA AGAATCGACG CGGTGCTCCA ACTTAAGGAG TCTGTAAGT AGCTGAAGGA
4601 TGAGGATATG GAGATCGACG ACGAGTTAGT ATGGATCCAT CCGGACAGTT GCCTGAAGGG AAGAAAGGGA TTCAGTACTA CAAAAGGAAA GTTGTATTCC
4701 TACTTTGAAG GCACCAAAAT CCATCAAGCA GCAAAAGATA TGGCGGAGAT AAAGTGCTCTG TTCCCAATG ACCAGGAAA CAACGAACAA CTGTGTGCCT
4801 ACATATTGGG GGAGACCATG GAAGCAATCC GCGAAAAATG CCCGGTCGAC CACAACCCGT CGTCTAGCCC GCCAAAAACG CTGCCGTGCC TCTGTATGTA
4901 TGCCATGACG CCAGAAAGGG TCCACAGACT CAGAAGCAAT AACGTCAAAG AAGTTACAGT ATGCTCCTCC ACCCCCCCTC CAAAGTACAA AATCAAGAAT
5001 GTTCAGAAGG TTCAGTGCAC AAAAGTAGTC CTGTTAACC CGCATACCCG CGCATTCGTT CCCGCCCGTA AGTACATAGA AGCACCAGAA CAGCCTGCAG
5101 CTCGCCCTGC ACAGGCCGAG GAGGCCCCCG GAGTTGTAGC GACACCAACA CCACCTGCAG CTGATAACAC CTCGCTTGAT GTCACGGACA TCTCACTGGA
5201 CATGGAAGAC AGTAGCGAAG GCTCACTCTT TTCGAGCTTT AGCGGATCGG ACAACTACCG AAGGCAGGTG GTGGTGGCTG ACGTCCATGC CGTCCAAGAG
5301 CCTGCCCTG TTCCACCGCC AAGGCTAAAG AAGATGGCCC GCCTGGCAGC GGCAAGAATG CAGGAAGAGC CAACTCCACC GGCAAGCACC AGCTCTGCCG
5401 ACGAGTCCCT TCACCTTCT TTTGATGGG TATCTATATC CTTCCGATCC CTTTTGACG GAGAGATGGC CCGCTTGGA CCGGCACAAC CCCCAGCAAG
5501 TACATGCCCT ACGGATGTGC CTATGCTTT CGGATCGTTT TCCGACGGAG AGATTGAGGA GTTGAGCCGC AGAGTAACCG AGTCGGAGCC CGTCTGTTT
5601 GGGTCATTTG AACCGGGCGA AGTGAATCA ATTATATCGT CCCGATCAGC CGTATCTTTT CCACCACGCA AGCAGAGACG TAGACGCAGG AGCAGGAGGA
5701 CCGAATACTG TCTAACCGGG GTAGGTGGGT ACATATTTTC GACGGACACA GGCCCTGGGC ACTTGCAAAA GAAGTCCGTT CTGCAGAAC AGCTTACAGA
5801 ACCGACCTTG GAGCGCAATG TTCTGGAAG AATCTACGCC CCGGTGCTCG ACACGTGCAAG AGAGGAACAG CTCAAAATCA GGTACCAGAT GATGCCACCC
5901 GAAGCCAACA AAAGCAGGTA CCACTCTCGA AAAGTAGAAA ACCAGAAAGC CATAACCACT GAGCGACTGC TTTCAGGGCT ACGACTGTAT AACTCTGCCA
6001 CAGATCAGCC AGAATGCTAT AAGATCACCT ACCCGAAACC ATCGTATTCC AGCAGTGTAC CAGCGAACTA CTCTGACCCA AAGTTTCTG TAGTGTGTTG
6101 TAACAATAT CTGCATGAGA ATTACCCGAC GGTAGCATCT TATCAGATCA CCGACGAGTA CGATGCTTAC TTGGATATGG TAGACGGGAC AGTCGCTTGC
6201 CTAGATACTG CAACTTTTTG CCCCAGCAAG CTTAGAAGTT ACCCGAAAAG ACACGAGTAT AGAGCCCCAA ACATCCGCG TCGGTTTCCA TCAGCGATGC
6301 AGAACACGTT GCAAAACGTC CTCATTGCCG CGACTAAAAG AAATGCAAC GTCACACAAA TCGTGAAGT GCCAACACTG GACTCAGCGA CATTCAACGT
6401 TGAATGCTTT CGAAAAATG CATGCAATGA CGAGTATTGG GAGGAGTTG CCCGAAAGCC AATTAGGATC ACTACTGAGT TCGTTACCGC ATACGTGGCC
6501 AGACTGAAAG GCCCTAAGGC CGCCGCACTG TTCGCAAGA CGCATAATTT GGTCCCATG CAAGAAGTGC CTATGGATAG ATTCTCATG GACATGAAAA
6601 GAGACGTGAA AGTTACACCT GGCACGAAAC ACACAGAAGA AAGACCGAAA GTACAAGTGA TACAAGCCGC AGAACCCCTG GCGACCGCTT ACCTATGCGG
6701 GATCCACCGG GAGTTAGTGC GCAGGCTTAC AGCCGTTTTG CTACCAACA TTCACACGCT CTTGACATG TCGGCGGAGG ACTTTGATGC AATCATAGCA
6801 GAACACTTCA AGCAAGGTGA CCCGTAAGT GAGACGATA TCGCTCGTT CGACAAAAGC CAAGACGACG CTATGGCGTT AACCGGCTG ATGATCTTGG
6901 AAGACCTGGG TGTGGACAA CCACTACTCG ACTTGATCGA GTGCGCTTTT GGAGAAATAT CATCCACCCA TCTGCCACG GGTACCCGTT TCAAAATCGG
7001 GCGGATGATG AAATCCGGAA TGTCTCTAC GCTCTTTGTC AACACAGTTC TGAATGTCT TATCGCCAGC AGAGTATTGG AGGAGCGGCT TAAACGTCC
7101 AAATGTGACG CATTATCGG CGACGACAAC ATTATACAG GAGTAGTATC TGACAAAGAA ATGGCTGAGA GGTGTGCCAC CTGGCTCAAC ATGGAGGTTA
7201 AGATCATTGA CGCAGTCATC GCGGAGAGAC CACCTTACTT CTGCGGTGGA TTCATCTTGC AAGATTCCGT TACCTCCACA GCGTGTCCG TGGCGGACCC
7301 CTTGAAAAGG CTGTTAAAGT TGGTAAACC GCTCCAGCC GACGATGAGC AAGACGAAGA CAGAAGACGC GCTCTGCTAG ATGAAACAAA GCGTGTGTTT
7401 AGAGTAGGTA TAACAGACAC CTTAGCAGTG GCCGTGGCAA CTCGGTATGA GGTAGACAAC ATCACACCTG TCCTGCTGGC ATTGAGAACT TTTGCCCAGA
7501 GCAAAAGAGC ATTTCAAGCC ATCAGAGGGG AAATAAGCA TCTCTACGGT GTCCTAAAT AGTCAGCATA GTACATTCA TCTGACTAAT ACCACAACAC
7601 CACCACCATG AATAGAGGAT TCTTAACAT GCTCGGCCG CGCCCTTCC CAGCCCCAC TGCCATGTGG AGGCCGCGGA GAAGGAGGCA GGGGGCCCCG
7701 ATGCTGCCCC GCAATGGGCT GGCTTCCCA ATCCAGCAAC TGACCACAGC CGTCAGTGCC CTAGTCATTG GACAGGCAAC TAGACCTCAA ACCCCACGCC
7801 CACGCCCGCC CGCCGCCAG AAGAAGCAGG CGCAAAGCA ACCACCGAAG CCGAAGAAAC CAAAAACACA GGAGAAGAAG AAGAAGCAAC CTGCAAAACC

Fig. 1B

7901 CAAACCCGGA AAGAGACAGC GTATGGCACT TAAGTTGGAG GCCGACAGAC TGTTCGACGT CAAAAATGAG GACGGAGATG TCATCGGGCA CGCACTGGCC
8001 ATGGAAGGAA AGGTAATGAA ACCACTCCAC GTGAAAGGAA CTATTGACCA CCCTGTGCTA TCAAAGCTCA AATTCACCAA GTCGTCAGCA TACGACATGG
8101 AGTTCGCACA GTTCCCGGTC AACATGAGAA GTGAGGCGTT CACCTACACC AGTGAACACC CTGAAGGGTT CTACAACTGG CACCACGGAG CGGTGCAGTA
8201 TAGTGGAGGC AGATTTACCA TCCCCCGCG AGTAGGAGGC AGAGGAGACA GTGGTCGTCC GATTATGGAT AACTCAGGCC GGGTTGTGCG GATAGTCCTC
8301 GGAGGGGCTG ATGAGGGAAC AAGAACC GCCC CTTTCGGTCG TCACCTGGAA TAGCAAAGGG AAGACAATCA AGACAACCCC GGAAGGGACA GAAGAGTGGT
8401 CTGCTGCACC ACTGGTCACG GCCATGTGCT TGCTTGGAAA CGTGAGCTTC CCATGCAATC GCCCGCCAC ATGCTACACC CGCGAACCAT CCAGAGCTCT
8501 CGACATCTC GAAGAGAAGC TGAACCAGA GGCCTACGAC ACCCTGCTCA ACGCCATATT GCGGTGCGGA TCGTCCGGA GAAGTAAAG AAGCGTCACT
8601 GACGACTTTA CTTGACCAG CCCGACTTG GGCACATGCT CGTACTGTCA CCATCTGAA CCGTGCTTTA GCCCGATTAA GATCGAGCAG GTCTGGGATG
8701 AAGCGGACGA CAACACCATA CGCATAAGA CTTCCGCCCA GTTTGGATAC GACCAAAGCG GAGCAGCAAG CTCAAATAAG TACCCTACA TGTCGCTCGA
8801 GCAGGATCAT ACTGTCAAAG AAGGCACCAT GGATGACATC AAGATCAGCA CCTCAGGACC GTGTAGAAG CTTAGCTACA AAGGATACTT TCTCCTCGCG
8901 AAGTGTCTC CAGGGGACAG CGTAACGGTT AGCATAGCGA GTAGCAACTC AGCAACGTCA TGCACAATGG CCCGCAAGAT AAAACCAAAA TTCGTGGGAC
9001 GGGAAAAATA TGACCTACCT CCCGTTACG GTAAGAAGAT TCCTTGACA GTGTACGACC GTCTGAAAGA AACCAACGCC GGCTACATCA CTATGCACAG
9101 GCCGGGACCG CATGCTATA CATCTATCT GGAGGAATCA TCAGGAAAG TTAACGCGAA GCCACCATCC GGAAGAACA TTACGTACGA GTGCAAGTGC
9201 GCGGATTACA AGACCGGAAC CGTTACGACC CGTACCGAAA TCACGGGCTG CACCGCCATC AAGCAGTGCG TCGCTATAA GAGCGACCAA ACGAAGTGGG
9301 TCTTCAACTC GCCGGACTCG ATCAGACAGC CCGACACAC GCCCAAGGG AAATTGCATT TGCCTTTCAA GCTGATCCCG AGTACCTGCA TGTCCCTGT
9401 TGCCACGCG CCGAACGTAG TACACGGCTT TAAACATC AGCCTCCAAT TAGACACAGA CCATCTGACA TTGCTACCA CCAGGAGACT AGGGGCAAAAC
9501 CCGGAACCAA CCACTGAATG GATCATCGGA AACACGGTTA GAAATTCAC CGTCGACCGA GATGGCCTGG AATACATATG GGGCAATCAC GAACCACTAA
9601 GGGTCTATG CCAAGAGTCT GCACCAGGAG ACCCTCACGG ATGGCCACAG GAAATAGTAC AGCATTACTA TCATCGCCAT CCGTGTGACA CCATCTTAGC
9701 CGTCGCATCA GCTGCTGTGG CGATGATGAT TGGCGTAACT GTTGACGAT TATGTGCTG TAAAGCGCGC CGTGAGTGCC TGACGCCATA TGCCCTGGCC
9801 CCAAATGCCG TGATTCCAA TTCGCTGGA CTTTGTGCT GTGTAGGTC GGCTAATGCT GAAACATTCA CCGAGACCAT GAGTTACTTA TGGTCGAACA
9901 GCCAGCGCTT CTTCTGGTC CAGCTGTGTA TACCTCTGCC CGCTGTGCT GTTCTAATGC GCTGTGCTC ATGTGCTCG CTTTTTTTG TGGTGGCCG
10001 CGCTACCTG GCGAAGGTAG ACGCTACGA ACATGCGACC ACTGTTCCAA ATGTGCCACA GATACCGTAT AAGGCACTTG TTGAAAGGGC AGGGTACGGC
10101 CCGCTCAATT TGGAGATTAC TGTCATGTCC TCGGAGGTTT TGCCTTCAC CAACCAAGAG TACATTACCT GCAAATTCAC CACTGTGTC CCCTCCCTA
10201 AAGTCAGATG CTGCGGCTCC TTGGAATGTC AGCCCGCCG TCACGACAG TATACCTGCA AGGTCTTTGG AGGGGTGTAC CCCTTCATGT GGGGAGGAGC
10301 ACAATGTTTT TCGACAGTG AGAACAGCCA GATGAGTGAG GCGTACGTC AATTGTCACT AGATTGCGCG ACTGACCAGC CGCAGGCGAT TAAGTGTGAT
10401 ACTGCCCGCA TGAAGTAGG ACTGCGTATA GTGTACGGA AACTACAGG TTTCTAGAT GTGTACGTA ACGGAGTCAC ACCAGGAACG TCTAAAGACC
10501 TGAAAGTCAT AGCTGGACCA ATTCAGCAT TGTTACACC ATTCGATCAC AAGGTCGTTA TCAATCGCGG CCGGTGTAC AACTATGACT TTCCGGAATA
10601 CGGAGCGATG AAACCAGGAG CGTTGGAGA CATTCAAGCT ACCTCCTGA CTAGCAAAGA CCTCATCGCC AGCACAGACA TTAGGCTACT CAAGCCTTCC
10701 GCCAAGAAGC TGCATGTCCC GTACACGAG GCCGATCTG GATTGAGAT GTGAAAAAC AACTCAGGCC GCCCACTGCA GGAACCCGCC CTTTTTGGT
10801 GCAAGATTGC AGTCAATCCG CTTGAGCGG TGGACTGCTC ATACGGGAAC ATTCCATTT CTATTGACAT CCCGAACGCT GCCTTTATCA GGACATCAGA
10901 TGACCACTG GTCTCAACAG TCAAATGTGA TGTCAGTGAG TGCACTTATT CAGCGGACTT CGGAGGGATG GCTACCTGC AGTATGTATC CGACCGCGAA
11001 GGACAATGCC CTGTACATTC GCATTCGAGC ACAGCAACCC TCCAAGATC GACAGTTCAT GTCTGGAGA AAGGAGCGGT GACAGTACAC TTCAGCACCG
11101 CGAGCCACA GCGCAACTTC ATTTATCGC TGTGTGGTAA GAAGACAACA TGCAATGCAG AATGCAAAACC ACCAGCTGAT CATATCGTGA GCACCCCGCA
11201 CAAAAATGAC CAAGAATTCC AAGCCGCCAT CTCAAAACT TCATGGAGTT GGCTGTTGC CTTTTTCGGC GCGCCTCGT CGCTATTAAT TATAGGACTT
11301 ATGATTTTTG CTTGACGAT GATGCTGACT AGCACACGAA GATGACCGT ACGCCCAAT GACCCGACCA GCAAACTCG ATGTACTCC GAGGAACTGA
11401 TGTGCATAAT GCATCAGGCT GGTATATTAG ATCCCGGCTT ACCCGGGGCA ATATAGCAAC ACCAAAACTC GACGTATTTT CGAGGAAGCG CAGTGCTATA
11501 TGCTGCGCAG TGTGCGCAA TAATCACTAT ATTAACCATT TATTCAGCGG ACGCCAAAA TCAATGTATT TCTGAGGAAG CATGGTGCAT AATGCCATGC
11601 AGCGTCTGCA TAACTTTTTA TTATTTCTTT TATTAATCAA CAAAATTTTG TTTTAAACAT TTC

Fig. 1c

S.A.AR86

A. Amino Acid Sequence of the Nonstructural Polyprotein

```

1      MEKPVVNVDP DQSPFFVQL QKSFPQFEV AQQVTPNDHA NARAFSHLAS KLIELEVPTT ATILDIGSAP ARRMFSEHQY HCVCPMRSPD DPDRMMKYAS
101    KLAEKACKIT NKNLHEKIKD LRTVLDTPDA ETPSLCFHND VTCNTRAEYS VMQDVYINAP GTIYHOAMKG VRTLYWIGFD TTQFMFSAMA GSYPAYNTNW
201    ADEKVLERN IGLCSTKLSE GRTGKLSIMR KKEKLPGRSV YFSVGSTLYP EHRASLQSWH LPSVFHLKGG QSYTCRCDTV VSCEGYVVKK ITISPGITGE
301    TVGYAVTNS EGFLLCKVTD TVKGERVSFP VCTYPATIC DQMTGIMATD ISPDDAQKLL VGLNQRIVIN GKTNRNTNTM QNYLLPIAQ GFSKWAKERK
401    EDLDNEKMLG TREKRLTYGC LWFARTKKVH SFYRPPGTQT IVKVPASFA FPMSSVWTTT LPMSLRQMKM LALQPKKEEK LLQVPEELVM EAKAAFEDAQ
501    EESRAEKLRE ALPPLVADKG IEAAAEVCE VEGLOADTGA ALVETPRGHV RIIPQANDRM IGQYIVVSPV SVLKNAKLAP AHPLADQVKI ITHSGRSGRY
601    AVEPYDAKVL MPAGSAVPWP EFLALSESAT LVYNEREFVN RKLHYIAMHG PAKNTEEEQY KYTKAELAET EYVFDVKKR CVKKEEASGL VLSGELTNPP
701    YHELALGLK TRPAVPYKVE TIGVIGTPGS GKSAIKSTV TARDLVTSKG KENCREEAD VLRLRGMQIT SKTVDSVMLN GCHKAVEVLY VDEAFRCHAG
801    ALLALIAIVR PRKKVVLCDG PKQCGFFNMN QLKVFHNPHE KDICTKTFYK FISRRCTQPV TAIVSTLHYD GKMKTTNPCK KNEIDITGA TKPKPGDIL
901    TCFRGWVKQL QIDYPGHEVM TAAASQGLTR KGVYAVROKV NENPLYAITS EHVNVLLTRT EDRLVWKTLO GDPWIKQLTN VPKGNFQATI EDWEAEHKG
1001   IAAINSPAPR TNPFSCCTNV CWAKALEPIL ATAGIVLTGC QWSELFPOFA DDKPHSAIYA LDVICIKFFG MDLTSGLFSK QSIPLTYHPA DSARPVAVHD
1101   NSPTGRKYGY DHAVAAELSR RFPVFQLAGK GTQLDLQGTGR TRVISAQHNH VPVNRNLPHV LPVEHKEKQP GPVEKFLSQF KHHSVLVISE KKIEAPHKRI
1201   EWIAPIGIAG ADKNYNLAFG FPPQARYDLV FINIGTKYRN HHFQQCEDHA ATLKTLRSRSLNCLNPGGTL VVKS YGYADR NSEDDVVTALA RKFVRVSAAR
1301   PECVSNTEM YLIFRQLDNS RTRQFTPHHL NCVISSVYEG TRDGVGAAPS YRTKRENIAD CQEEAVVNAA NPLGRPGEGV CRAIYKRWPV SFTDSATETG
1401   TAKLTYCOGK KVIHAGVPDF RKHPEAEALK LLQNAHYHVA DLVNEHNKIS VAIPLLSTGI YAAGKDRLEV SLNCLTTALD RTDADVITYC LDKKWKERID
1501   AVLQKLESVT ELKDEDEMEID DELVWIHPDS CLKGRKGFST TKGKLYSYFE GTKFHQAAKD MAEIKVLFPN DOESNEQLCA YILGETMEAI REKCPVDHNP
1601   SSSPKTLPC LCMYAMTPER VHLRSLNNVK ETVCSSTPL PKYKIKNVQK VQCTKVVLFN PHTPAFVPAR KYIEAPEQA APPAQAEAP GVVATTPTPA
1701   ADNTSLDVTI ISLDMEDSSE GSLFSSFGS DNYRQVVA DVHAVQEAP VPPRLKKMA RLAAARMQEE PTPASTSSA DESLHLSFDG VSISFGSLFD
1801   GEMARLAAAQ PPASTCPTDV PMSFGSFDG EIEELSRVT ESEFVLGSGF EPGEVNISI SRSVSPFRP KQRRRRRRR TEVCLTGVGG YIFSTDTGPG
1901   HLQKKSVLQN QLTEPTLERN VLERIYAPV DTSKEEQKL RYQMMPTAN KSRYSRKYVE NOKAITERL LSLRLVNSA TDQPECYKTI YPKPSYSSV
2001   PANYSDPKFA VAVCNLYLHE NYPTVASYQI TDEYDAYLDM VDGTVACLDI ATFCPAKLRS YPKRHEYRAP NRSVAVPSAM QNTLQNVLIA ATKRCNVTQ
2101   MRELPTLDSA TFNVECFRKY ACNDEYWEFF ARKPIRITTE FVTAYVARLK GPKAAALFAK THNLVPLQEV PMDRFVMDMK RDVKTVPGTK HTEERPKVQV
2201   IQAAEPLATA YLCGIHREL VRLTAVLLPN IHTLFDMSAE DFDIAIEHF KOGDPVLETD IASFDKSQDD AMALTGLML EDLGVDQPLL DLIECAFGEI
2301   STHLPTGTR FKFGAMMKSG MFLTLFVNTV LNVVIASRVL EERLTKSKCA AFIGDDMIH GVVSDKEMAE RCATWLNMEV KIIDAVIGER PPYFCGGFTL
2401   QDSVTSTACR VADPLKRLFK LGKPLPADDE QDEDRRRALL DETKAWFRVG ITDLAVAVA TRYEVNNTT VLLALRTFAQ SKRAFQAIRG EIKHLYGGFK

```

B. Amino Acid Sequence of the Structural Polyprotein

```

1      MNRGFFNMLG RRPFPAPTAM WRPRRRRQAA PMPARNGLAS QIQQLTTAVS ALVIGQATRP QTPRPRPPPR QKKQAPKQPP KPKPKTQEK KKKQPAKPKP
101    GKRQRMALKL EADRLFDVKN EDGDIVGHAL AMEGKVMKPL HYKGTIDHPV LSKLKFTKSS AYDMEFAQLP VNMREAFYI TSEHPGEFYN WHHGAVQYSG
201    GRFTIPRGV GRGDSGRPIM DNSGRVVAIV LGGADEGTRT ALSVVTWNSK GKTIKTTPG TEWSAAPLV TAMCLLGNVS FPCNRPPTCY TREPSRALDI
301    LEENVNHEAY DTLNLAILRC GSSGRSKRSV TDDFTLTSPY LGTCSYCHHT EPCFSPKIE QVWDEADDNT IRIQTSAQFG YDQSGAASN KYRYSLEQD
401    HTVKEGTMDI IKISTGPRC RLSYKGYFLL AKCPGDSVT VSIASSNSAT SCTMARKKP KFGREKYDL PPVHGKIPC TVYDRLKETT AGYITMHRPG
501    PHAYTSYLEE SSGKYVAKPP SGKNITYECK CGDYKTGTVT TRTEITGCTA IKQCVAYKSD QTKWVFNSPD SIRHADHTAQ GKLHLFPKLI PSTCMVPAH
601    APNVVHGFKH ISLQLDTHL TLLTTRRLGA NPEPTTEWII GNTVRNFTVD RDGLEIYWGN HEPVRVYAE SAPGDPHGWP HEIVQHYHR HPVYITLAVA
701    SAAVAMMIGV TVAALCACKA RRECLTPYAL APNAVITSL ALLCCVRSAN AETFTETMSY LWSNSQFFW VQLCIPLAAV VVLMRCCSCC LPFLVVGAY
801    LAKVDAYEHA TTVPNVQIP YKALVERAGY APLNLEITVM SSEVLPTNQ EYITCKFTTV VPSPKVRCCG SLECPAAHA DYTCKVFGGV YPFMWGGAQC
901    FCDSENSQMS EAYVELSDC ATDHAQIKV HTAAMKVGLR IVYGNNTSFL DYYVNGVTPG TSKDLKVIAG PISALFTPD HKVVINRGLV YNYDFPEYGA
1001   MKPGAFGDIQ ATSLTSKDLI ASTDIRLLKP SAKNVHVPYT QAASGFEMWK NNSGRPLQET APFGCKIAVN PLRAVDCSYG NIPISIDFN AAFIRTSAP
1101   LVSTVKCDVS ECTYSADFGG MATLQYVSDR EGQCPVHSHS STATLQESTV HVLEKGAVTV HFSTASPOAN FIVSLCGKKT TCNAECKPPA DHIVSTPHKN
1201   DQEFQAAISK TSWSWLFALE GGASSLLIG LMIFACSMML TSTRR

```

FIG. 2

Nucleotide Sequence of Girdwood S.A.

1 NTTGNCGGCG TAGTATACAC TATTGAATCA AACAGCCGAC CAATTGCACT ACCATCACA TGGAGAAGCC AGTAGTTAAC GTAGACGTAG ACCCGCAGAG
101 TCCGTTTGTG GTGCAACTGC AAAAGAGCTT CCCGCAATTT GAGGTAGTAG CACAGCAGGT CACTCCAAAT GACCATGCTA ATGCCAGAGC ATTTTCGCAT
201 CTGGCCAGTA AACTAATCGA GCTGGAGGTT CCTACCACAG CGACGATTTT GGACATAGGC AGCGCACC GG CTCGTAGAAT GTTTTCCGAG CACCACTACC
301 ATTGCGTTTG CCCATGCGT AGTCCAGAAG ACCCGGACCG CATGATGAAA TATGCCAGCA AACTGGCGGA AAAAGCATGC AAGATTACGA ATAAGAACTT
401 GCATGAGAAG ATCAAGGACC TCCGGACCGT ACTTGATACA CCGGATGCTG AAACGCCATC ACTCTGCTTC CACAACGATG TTACTTGCAA CACGCGTGCC
501 GAGTACTCCG TCATGCAGGA CGTGATACAT AACGCTCCCG GAACTATTTA CCATCAGGCT ATGAAAGGCG TCGGACCCCT GTACTGGATT GGCTTCGATA
601 CCACCCAGTT CATGTTCTCG GCTATGGCAG GTTCGTACCC TCGGTACAAC ACCAACTGGG CCGACGAAAA AGTCTCTGAA GCGCGTAACA TCGGACTCTG
701 CAGCACAAAG CTGAGTGAAG GCAGGACAGG AAAGTTGTCTG ATAATGAGGA AGAAGGAGTT GAAGCCCGGG TCACGGGTTT ATTTCTCCGT TGGATCGACA
801 CTTTACCCAG AACACAGAGC CAGCTTGCG AGCTGGCATC TTCCATCGGT GTTCCACCTG AAAGGAAAGC AGTCGTACAC TTCCCGCTGT GATACAGTGG
901 TGAGCTGCGA AGGCTACGTA GTGAAGAAAA TCACCATCAG TCCCGGGATC ACGGGAGAAA CCGTGGGATA CCGCGTTACA AACAATAGCG AGGGCTTCTT
1001 GCTATGCAAA GTTACCGATA CAGTAAAAGG AGAACGGGTA TCGTTCCCGG TGTGCACGTA TATCCCGGCC ACCATATGCG ATCAGATGAC CGGCATAATG
1101 GCCACGGATA TCTCACCTGA CGATGCACAA AAATTTCTGG TTGGGCTCAA CCAGCGAATC GTCATTAACG GTAAGACTAA CAGGAACACC AATACCATGC
1201 AAAATTACCT TGTGCCAATC ATTGCACAAG GGTTCAGCAA ATGGGCCAAG GAGCGCAAAG AAGACCTTGA CAATGAAAAA ATGCTGGGTA CCAGAGAGCG
1301 CAAGCTTACA TATGGCTGCT TGTGGGCGTT TCGCACTAAG AAAGTGCACT CGTTCTATCG CCCACCTGGA ACGCAGACCA TCGTAAAAGT CCCAGCCTCT
1401 TTTAGCGCTT TCCCATGTC ATCCGTATGG ACTACCTCTT TGCCCATGTC GCTGAGGCGA AAGATAAAAT TGGCATTACA ACCAAAAGAG GAGGAAAAAC
1501 TGCTGCAAGT CCCGGAGGAA TTAGTCATGG AGGCCAAGGC TGCTTTCGAG GATGCTCAGG AGGAATCCAG AGCGGAGAAG CTCGAGAAG CACTCCACC
1601 ATTAGTGCA GACAAAGGTA TCGAGGCAGC CGCGGAAGTT GTCTGCGAAG TGGAGGGGCT CCAGGCGGAC ATCGGAGCAG CACTCGTCGA AACCCCGCG
1701 GGTCAATGTA GGATAATACC ACAAGCAAAT GACCGTATGA TCGGACAGTA CATCGTTGTC TCGCCAACCT CTGTGCTGAA GAACGCTAAA CTCGCACCA
1801 CACACCCGCT AGCAGACCAG GTTAAGATCA TAACGCACCT CGGAAGATCA GGAAGGTATG CAGTCGAACC ATACGACGCT AAAGTACTGA TGCCAGCAGG
1901 AAGTGCCGTA CCATGGCCAG AATTCTTAGC ACTGAGTGAG AGCGCCACGC TAGTGTAACA CGAAAGAGAG TTTGTGAACC GCAAGCTGTA CCATATTGCC
2001 ATGCACGGTC CCGTAAGAA TACAGAAGAG GAGCAGTACA AGGTTACAAA GGCAGAGCTC GCAGAAACAG AGTACGTGTT TGACGTGGAC AAGAAGCGAT
2101 GCGTCAAGAA GGAAGAAGCC TCAGGACTTG TCCTCTCGGG AGAAGTGACC AACCCCGCCT ATCAGGAACT AGCTCTTGAG GGAAGTGAAG CTCGACCCGT
2201 GGTCCCGTAC AAGGTTGAAA CAATAGGAGT GATAGGCGCA CCAGGATCGG GCAAGTCGGC TATCATCAAG TCAACTGTCA CGGCACGTGA TCTTGTACC
2301 AGCGGAAAGA AAGAAAACTG CCGCGAAATT CAGGCCGATG TGCTACGGCT GAGGGGCATG CAGATCACGT CGAAGACAGT GGATTCCGTT ATGCTCAACG
2401 GATGCCGCAA AGCCGTAGAA GTGCTGTATG TTGACGAAGC GTTCGCGTGC CACGCAGGAG CACTACTTGC TTGATTGCA ATCGTCAGAC CCGTCATAA
2501 GGTAGTGCTA TCGCGAGACC CTAAGCAATG CGGATTCTTC AACATGATGC AACTAAAGGT ATATTTCAAC CACCCGGAAA AAGACATATG TACCAAGACA
2601 TTCTACAAGT TTATCTCCCG ACGTTGCACA CAGCCAGTCA CGGCTATTGT ATCGACACTG CATTACGATG GAAAAATGAA AACCAAAAAC CCGTGCAAGA
2701 AGAACATCGA AATCGACATT ACAGGGGCCA CGAAGCCGAA GCCAGGGGAC ATCATCTGA CATGCTTCCG CGGGTGGGTT AAGCAACTGC AAATCGACTA
2801 TCCCGGACAT GAGGTAATGA CAGCCGCGGC CTCACAAGGG CTAACCAGAA AAGGAGTATA TGCCGTCCGG CAAAAAGTCA ATGAAAAACC GCTGTACGG
2901 ATCACATCAG AGCATGTGAA CGTGCTGCTC ACCCGCACTG AGGACAGGCT AGTATGGAAG ACTTTACAGG GCGACCCATG GATTAAGCAG CTCACTAACG
3001 TACCAAAAGG AAATTTTCAA GCCACCATCG AGGACTGGGA AGCTGAACAC AAGGGAATAA TTGCTGCGAT AAACAGTCCC GCTCCCGTA CCAATCCGTT
3101 CAGCTGCAAG ACTAACGTTT GCTGGCGGAA ACGACTGGAA CCGATACTGG CCACGGCCGG TATCGTACTT ACCGGTTGCC AGTGGAGCGA GCTGTTECCA
3201 CAGTTTGCAG ATGACAAACC AACTCGGCC ATCTACGCCC TGGACGTAAT CTGCATTAAG TTTTTCGGCA TGGACTTGAC AAGCGGACTG TTTTCAAAC
3301 AGAGCATCCC GTTAACGTAC CATCTGCGG ATTCAGCGAG GCCAGTAGCT CATTGGGACA ACAGCCAGG AACCCGCAAG TAGGGTACG ATCAGCCGCT
3401 TGCCCGCGAA CTCTCCCGTA GATTTCGGT GTTCCAGCTA GCTGGGAAAG GCACACAGCT TGATTGCAAG ACGGGCAGAA CTAGAGTTAT CTCGACACAG
3501 CATAACTTGG TCCCACTGAA CCGCAATCTC CCGCAGCCT TAGTCCCCGA GCACAAGGAG AAACAACCCG GCCCGGTCAA AAAATTCTTG AGCCAATTCA
3601 AACACCACTC CGTACTTGTG GTCTCAGAGG AAAAAATTGA AGTCCCCAC AAGAGAATCG AATGGATCGC CCCGATTGGC ATAGCCGGCG CTGATAAGAA
3701 CTACAACCTG GCTTTCGGGT TTCCGCCGCA GGCACGGTAC GACCTGGTGT TTATCAATAT TGGAACTAAA TACAGAAACC ATCACTTTCA GCAGTGCGAA

Fig. 3A

3801 GACCATGCGG CGACCTTGAA AACCTCTCG CGTTCGGCCC TGAAGTGCCT TAACCCCGGA GGCACCCTCG TGGTGAAGTC CTACGGTTAC GCCGACCGCA
3901 ATAGTGAGGA CGTAGTCACC GCTCTTGCCA GAAAATTGT CAGAGTGTCT GCACGAGGC CAGAGTGCCT CTCAAGCAAT ACAGAAATGT ACCTGATCTT
4001 CCGACAATA GACAACAGCC GCACACGACA ATTCACCCCG CATCATCTGA ATTGTGTGAT TTCGTCCGTG TACGAGGGTA CAAGAGACGG AGTTGGAGCC
4101 GCACCGTCAT ACCGCACTAA AAGGGAGAAC ATTGCTGATT GTCAAGAGGA AGCAGTTGTC AATGCAGCCA ATCCGCTGGG CAGACCAGGC GAAGGAGTCT
4201 GCCGTGCCAT CTATAAACGT TGGCCGAACA GTTTCACCGA TTCAGCCACA GAGACCGGCA CCGCAAACT GACTGTGTGC CAAGGAAAGA AAGTGATCCA
4301 CGCGTTGGC CCTGATTTCC GGAACACCC AGAGGCAGAA GCCCTGAAAT TGCTGCAAAA CGCCTACCAT GCAGTGGCAG ACTTAGTAAA TGAACATAAT
4401 ATCAAGTCTG TCGCCATCCC ACTGCTATCT ACAGGCATTT ACGCAGCCGG AAAAGACCGC CTTGAAGTAT CACTTAACTG CTTGACAACC GCGTAGATA
4501 GAACTGATGC GGACGTAAAC ATCTACTGCC TGGATAAGAA GTGGAAGGAA AGAATCGACG CGGTGCTCCA ACTTAAGGAG TCTGTAATAG AGCTGAAGGA
4601 TGAGGATATG GAGATCGACG ACGAGTTAGT ATGGATCCAT CCGGACAGTT GCCTGAAGGG AAGAAAGGGA TTCAGTACTA CAAAAGGAAA GTTGTATTCG
4701 TACTTTGAAG GCACCAAAAT CCATCAAGCA GCAAAAGATA TGGCGGAGAT AAAGTCTCTG TTCCCAAAATG ACCAGGAAAG CAACGAGCAA CTGTGTGCCT
4801 ACATATTGGG GGAGACCATG GAAGCAATCC GCGAAAAATG CCCGGTCGAC CACAACCCGT CGTCTAGCCC GCCAAAAACG CTGCCGTGCC TCTGCATGTA
4901 TGCCATGACG CCAGAAAGGG TCCACAGACT CAGAAGCAAC AACGTCAAAG AAGTTACAGT ATGCTCCTCC ACCCCCTTC CAAAGTACAA AATCAAGAAC
5001 GTTCAGAAGG TTCAGTGCAC AAAAGTAGTC CTGTTTAAAC CGCATACCCC TGCATTCTGT CCCGCCGTA AGTACATAGA AGCGCCAGAA CAGCCTGCAG
5101 CTCGCTGAC ACAGGCCGAG GAGGCCCCG AAGTTGCAGC AACACCAACA CCACCTGCAG CTGATAACAC CTCGCTGAT GTCACGGACA TCTCACTGGA
5201 CATGGAAGAC AGTAGCGAAG GCTCACTCTT TTCGAGCTTT AGCGGATCGG ACAACTCTAT TACTAGTATG GACAGTTGGT CGTCAGGACC TAGTTCACTA
5301 GAGATAGTAG ACCGAAGGCA GGTGGTGGTG GCTGACGTCC ATGCCGTCCA AGAGCCTGCC CCGTCTCCAC CGCCAAGGT AAAGAAGATG GCCCGCTGG
5401 CAGCGGCAAG AATGCAGGAA GAGCCAATC CACCGGCAAG CACCAGCTCT CGGACGAGT CCCTTCACTT TTCTTTGGT GGGGTATCCA TGTCTTCGG
5501 ATCCCTTTTC GACGGAGAGA TGGCGCCCTT GGCAGCGGCA CAACCCCGG CAAGTACATG CCCTACGGAT GTGCCTATGT CTTTCGGATC GTTTTCGAC
5601 GGAGAGATTG AGGAGCTGAG CCGCAGAGTA ACCGAGTCTG AGCCCGTCTT GTTTGGGTCA TTGAACCGG GCGAAGTGAA CTCAATTATA TCGTCCCGAT
5701 CAGTTGTATC TTTTCCACCA CGCAAGCAGA GACGTAGACG CAGGAGCAGG AGGACCGAAT ACTGACTAAC CGGGGTAGGT GGTACATAT TTTGACGGA
5801 CACAGGCCCT GGGCACTTGC AAATGGAGTC CGTCTGCAG AATCAGCTTA CAGAACCAGC CTTGGAGCGC AATGTTCTGG AAAGAATCTA CGCCCGGTG
5901 CTCGACACGT CGAAGAGGA ACAGCTCAAA CTCAGGTACC AGATGATGCC CACCGAAGCC AACAAAAGCA GGTACCAGTC TAGAAAAGTA GAAAATCAGA
6001 AAGCCATAAC CACTGAGCGA CTGCTTTCAG GGCTACGACT GTATAACTCT GCCACAGATC AGCCAGAATG CTATAAGATC ACCTACCCGA AACCATCGTA
6101 TTCCAGCAGT GTACCGGCGA ACTACTCTGA CCAAAAGTTT GCTGTAGCTG TTGCAACAA CTATCTGCAT GAGAATTACC CGACGGTAGC ATCTTATCAG
6201 ATCACCGACG AGTACGATGC TTACTTGGAT ATGGTAGACG GGACAGTCCG TTGCCTAGAT ACTGCAACTT TTTGCCCGC CAAGCTTAGA AGTTACCCGA
6301 AAAGACACGA GTATAGAGCC CCAACACTC GCAGTGGGT TCCATCAGCG ATGCAGAACA CGTTGCAAAA CGTGCTCATT GCCCGACTA AAAGAACTG
6401 CAACGTCACA CAAATGCGTG AATTGCCAAC ACTGGACTCA GCGACATTC ACGTTGAATG CTTTCGAAAA TATGCATGTA ATGACGAGTA TTGGGAGGAG
6501 TTTGCCCGAA AGCCAATTAG GATCACTACT GAGTTCTGTA CCGCATACGT GGCCAGACTG AAAGGCCCTA AGGCCCGCG ACTGTTCCGA AAGACGCATA
6601 ATTTGGTCCC ATTGCAAGAA GTGCCTATGG ATAGGTTCGT CATGGACATG AAAAGAGACG TGAAAGTTAC ACCTGGCAGG AAACACACAG AAGAAAGACC
6701 GAAAGTACAA GTGCTACAAG CCGCAGAACC CTTGGCGACC GCTTACCTGT GCGGGATCCA CCGGGAATTA GTGCGCAGGC TTACAGCCGT CTTGCTACCC
6801 AACATTACA CGCTTTTGA CATGTCCGGG GAGGACTTTG ATGCAATCAT AGCAGAACAC TTCAAGCAAG GTGACCCGGT ACTGGAGACG GATATCGCCT
6901 CGTTGACAA AAGCCAAGAC GACGCTATGG CGTTAACTGG CCGTATGATC TTGGAAGACC TGGGTGTGGA CCAACCACTA CTCGACTTGA TCGAGTGGC
7001 CTTTGGAGAA ATATCATCCA CCCATCTGCC CACGGGTACC CGTTTCAAAT TCGGGCGCAT GATGAAATCC GGAATGTTCC TCACGCTCTT TGTCAACACA
7101 GTTCTGAATG TCGTTATCGC CAGCAGAGTA TTGGAGGAGC GGCTTAAAA GTCCAAATGT GCAGCATTTA TCGGCGACGA CAACATCATA CACGGAGTAG
7201 TATCTGACAA AGAAATGGCT GAGAGGTGTG CCACCTGGCT CAACATGGAG GTTAAGATCA TTGACGAGT CATCGGCGAG AGACCCGCTT ACTTCTGCGG
7301 TGGATTATC TTGCAAGATT CGGTACCTC CACAGCGTGT CGCGTGGCGG ACCCTTGAA AAGGCTGTTT AAGTTGGGTA AACCGCTCCC AGCCGACGAC
7401 GAGCAAGACG AAGACAGAAG ACGCGCTCTG CTAGATGAAA CAAAGGCGTG GTTTAGAGTA GGTATAACAG ACACCTTAGC AGTGGCCGTG GCAACTCGGT
7501 ATGAGGTAGA CAACATCACA CCGTCTCTGC TGGCATTGAG AACTTTTGGC CAGAGCAAAA GAGCATTTCA AGCCATCAGA GGGGAAATAA AGCATCTCTA
7601 CGGTGTCTCT AAATAGTCAG CATAGCACAT TTCATCTGAC TAATACCACA ACACCACCAC CATGAATAGA GGATTCTTTA ACATGCTCGG CCGCCGCCCC
7701 TTCCCGCCCC CCACTGCCAT GTGGAGGCCG CGGAGAAGGA GGCAGGCGGC CCCGATGCCT GCCCGCAATG GGCTGGCTTC CCAATCCAG CAACTGACCA
7801 CAGCCGTCAG TGCCCTAGTC ATTGGACAGG CAACTAGACC TCAAAACCCA CGCCACGCC CGCCGCGCG CAGAGGAAG CAGGCGCCAA AGCAACCACC

Fig. 3B

7901 GAAGCCGAAG AAACCAAAAA CACAGGAGAA GAAGAAGAAG CAACCTGCAA AACCCAAACC CGGAAAGAGA CAACGTATGG CACTCAAGTT GGAGGCCGAC
8001 AGACTGTTTCG ACGTCAAAAA TGAGGACGGA GATGTCATCG GGCACGCACT GGCCATGGAA GGAAAGGTAA TGAAACCACT CCACGTGAAA GGAACATTGG
8101 ACCACCCTGT GCTATCAAAAG CTCAAATCCA CCAAGTCGTC AGCATACGAC ATGGAGTTTCG CACAGTTGCC GGTCAACATG AGAAGTGAGG CGTTCACCTA
8201 CACCAGCGAA CACCCTGAAG GGTTTTACAA CTGGCACCAC GGAGCGGTGC AGTATAGTGG AGGTAGATTT ACCATCCCCC GCGGAGTAGG AGGCAGAGGA
8301 GACAGTGTGC GTCCGATTAT GGATAACTCA GGCCGGTTG TCGCGTAGT CCTCGGAGGG GCTGATGAGG GAACAAGAAC TGCCCTTTTCG GTCGTCACCT
8401 GGAATAGCAA AGGGAAGACA ATCAAGACAA CCCCAGGAGG GACAGAAGAG TGGTCTGCAG CACCCTGGT CACGGCCATG TGCTTGCTTG GAAACGTGAG
8501 CTTCCCATGC AATCGCCCCG CCACATGCTA CCCCCGGA CCATCCAGAG CTCTTGACAT CTTGAAGAG AACGTGAACC ACGAGGCCA CGACACCCTG
8601 CTCACGCCA TATTGCGTG CGGATCGTCC GGCAGAAGCA AAAGAAGCGT CACTGACGAC TTTACCTTGA CCAGCCCGTA CTTGGGCACA TGCTCGTACT
8701 GTCACCATAC TGAACCGTGC TTTAGCCCGA TTAAGATCGA GCAGGTCTGG GATGAAGCGG ACGACAACAC CATACGCATA CAGACTTCCG CCCAGTTTGG
8801 ATACGACCAA AGCGGAGCAG CAAGCTCAA TAAGTACCGC TACATGTCGC TCGAGCAGGA TCATACCGTC AAAGAAGGCA CTATGGATGA CATCAAGATC
8901 AGCACCTCAG GACCGTGTAG AAGGCTTAGC TACAAAGGAT ACTTTCTCT CGCGAAGTGT CCTCCAGGGG ACAGCGTAAC GGTTAGTATA GCGAGTAGCA
9001 ACTCAGCAAC GTCATGCACA ATGGCCCGCA AGATAAAACC AAAATTCGTG GGACGGGAAA AATATGACCT ACCTCCCGTT CACGGTAAGA AGATTCTTGG
9101 CACAGTGTAC GACCGTCTGA AAGAAACAAC CGCCGGCTAC ATCACTATGC ACAGGCCGGG ACCGCACGCC TATACGTCTT ATCTGGAGGA ATCATCAGGG
9201 AAAGTCTACG CGAAGCCACC ATCCGGAAAG AACATTACGT ACGAGTGCAA GTGCGCGCAT TACAAGACCG GTACCGTTAC GACCCGTACC GAAATCACGG
9301 GCTGCACCCG CATCAAGCAG TCGCTCGCCT ATAAGAGCGA CCAAACGAAG TGGGTCTTCA ATTCGCCGGA CTTGATCAGA CATGCCGACC ACACGGCCCA
9401 AGGGAATTTG CATTACCTT TCAAGCTGAT CCGGAGTACC TGCATGGTCC CTGTTGCCCA CGCGCCGAAC GTAGTACACG GCTTTAAACA CATCAGCCTC
9501 CAATTAGACA CAGACCACCT GACATTGCTC ACCACCAGGA GACTAGGGGC AAATCCGGAA CCAACTACTG AATGGATCAT CGGAAAGACG GTTAGAAACT
9601 TCACCGTCCA CCGAGATGGC CTGGAATACA TATGGGGCAA TCACGAACCG GTAAGGTCT ATGCCAAGA GTCTGCACCA GGAGACCCTC ACGGATGGCC
9701 ACACGAAATA GTACAGCATT ACTACCATCG CCATCCTGTG TACACCATCT TAGCCGTCGC ATCAGCTGCT GTGGCGATGA TGATTGGCT AACTGTTGCA
9801 GCATTATGTG CTGTAAAGC GCGCCGTGAG TGCCTGACGC CATATGCCCT GGCCCCAAAT GCGTGATTC CAATTCGCT GGCACTTTTCG TGCTGTGTTA
9901 GGTCGGCTAA TGCTGAAACA TTCACCGAGA CCATGAGTTA CCTATGGTGC AACAGCCAGC CATTCTTCTG GGTCCAGCTG TGTATACCCC TGGCCGCTGT
10001 CATCGTTCTA ATGCGCTGTT GCTCATGCTG CCTGCCTTTT TTAGTGTTG CCGGCGCTA CTTGGCGAAG GTAGACGCTC ACGAACATGC GACCACTGTT
10101 CCAATGTGTC CACAGATACC GTATAAGGA CTTGTTGAAA GGGCAGGGTA CGCCCCGCTC AATTGGGAGA TTAGTGTAT GTCTCGGAG GTTTTGCTT
10201 CCACCAACCA AGAGTACATC ACCTGCAAAAT TCACCACTGT GGTCCCTTCC CTAAGTCA AATGCTGCGG CTCCTTGGAA TGTCAGCCCC CCGCTCAGGC
10301 AGACTATACC TGCAAGGTCT TTGGAGGGGT GTACCCCTTC ATGTGGGGAG GAGCACAATG TTTTGGCAG AGTGAGAACA GCCAGATGAG TGAGGCGTAC
10401 GTCGAATTGT CAGCAGATTG CCGCACTGAC CACGCGCAGG CGATTAAGGT GCATACTGCC GCGATGAAAG TAGGACTACG TATAGTGAC GGAACACTA
10501 CCAATTCTCT AGATGTGTAC GTGAACGGAG TCACACCAGG AACGTCTAAA GACCTGAAAG TCATAGCTGG ACCAATTTCA GCATCGTTTA CACCATTCGA
10601 TCACAAGGTC GTTATCCATC CGGGCTGCT GTACAACTAT GACTTCCCGG AATACGGAGC GATGAAACCA GGAGCGTTTG GAGACATTCA AGCTACCTCC
10701 TTGACTAGCA AAGATCTCAT CGCCAGCACA GACATTAGAC TACTCAAGCC TTCCGCGAAG AACGTGCATG TCCCGTACAC GCAGGCCGCA TCTGGATTGG
10801 AGATGTGGAA AAACAACCTA GGCCGCCAC TGCAGGAAAC CGCCCTTTC GGTGCAAGA TTGCAGTCAA TCCGCTTCA GCGGTGGAAT GCTCATACGG
10901 GAACATTCCC ATCTCTATCG ACATCCGAA CGCTGCCTTT ATCAGGACAT CAGATGCACC ACTGGTCTCA ACAGTCAAAAT GTGATGTCAG TGAGTCACT
11001 TACTCAGCGG ACTTCGGCGG GATGGCTACC CTGCAGTATG TATCCGACCG CGAAGGACAA TGCCCTGTAC ATTCGCAATC GAGCACAGCA ACCCTCCAAG
11101 AGTCGACAGT TCATGTCTG GAGAAAGGAG CGGTGACAGT ACACCTCAGC ACCGCGAGCC CACAGGCGAA CTTTATTGTA TCGCTGTGTG GTAAGAAGAC
11201 AACATGCAAT GCAGAATGCA AACCAACAGC TGACCATATC GTGAGCACCC CGCACAAAAA TGACCAAGAA TTCCAAGCCG CCATCTCAAA AACTTCATGG
11301 AGTTGGCTGT TTGCCCTTTT CGGCGGCGCC TCGTCGCTAT TAATTATAGG ACTTATGATT TTTGCTTGCA GCATGATGCT GACTAGCACA CGAAGATGAC
11401 CGCTACGCCC CAATGACCCG ACCAGCAAAA CTCGATGTAC TTCCGAGGAA CTGATGTGCA TAATGCATCA GGCTGGTATA TTAGATCCCC GCTTACCGCG
11501 GGCAATATAG CAACACCAAA ACTCGACGTA TTTCCGAGGA AGCGCAGTGC ATAATGCTGC GCAGTGTTGC CAAATAATCA CTATATTAAC CATTATTATTA
11601 GCGGACGCCA AAACCTAATG TATTTCTGAG GAAGCATGGT GCATAATGCC ATGCAGCGTC TGCATAACTT TTTATTATTT CTTTATTATA TCAACAAAAAT
11701 TTTGTTTTTA ACATTTN

Fig. 3c

8/12

Girdwood S.A.

A. Amino Acid Sequence of the NonStructural Polyprotein

```

1      MEKPVVNDV DPQSPFVVQL QKSFPQFEVV AQQVTPNDHA NARAFSHLAS KLIELEVPTT ATILDIGSAP ARRMFSEHQY HCVCPMRSPE DPDRMMKYAS
101    KLAEKACKIT NKNLHEKIKD LRTVLDTPDA ETPSLCFHND VTCNTRAEYS VMQDVYINAP GTTYHQAMKG VRTLYWIGFD TTQFMFSAMA GSYPAYNTNW
201    ADEKVLEARN IGLCSTKLE GRTGKLSMR KKELKPGSRV YFSVGSTLYP EHRSLSQSWH LPSVFHLKKG QSYTCRCDTV VSCGEYVVKK ITSPGITGE
301    TVGYAVTNS EGFLLCKVTD TVKGERVSFP VCTYPATIC DQMTGIMATD ISPDDAQKLL VGLNQIRVIN GKTNRNTNTM QNYLLPIAQ GFSKWAKERK
401    EDLDNEKMLG TRERKLTGCG LWAFRTKKVH SFYRPPGTQT IVKVPASFA FPMSSVWTTT LPMSLRQKIK LALQPKKEEK LLQVPEELVM EAKAAFEFAQ
501    EESRAEKLRE ALPPLVADKG IEAAAEVYCE VEGLOQADIGA ALVETPRGHV RIIPQANDRM IGQYTVVSPT SVLKNAKLAP AHPLADQVKI ITHSGRSGRY
601    AVEPYDAKVL MPAGSAVPWP EFLALSESAT LVYNEREFVN RKLYHIAMHG PAKNTEEEQY KVTKAELAET EYVFDVDKKR CVKKEEASGL VLSGELTNPP
701    YHELALIEGLK TRPVVPYKVE TIGVIGAPGS GKSAIKSTV TARDLVTSGK KENCREIQAD VLRLRGMQIT SKTVDSVMLN GCRKAVEVLY VDEAFACHAG
801    ALLALIAIVR PRHKVVLGCD PKQCGFFNMM QLKVYFNHPE KDICTKTFYK FISRRCTQPV TAIVSTLHYD GKMKTTPNCK KNIEIDITGA TKPKPGDIIL
901    TCFRGWVYQL QIDYPGHEVM TAAASQGLTR KGVYAVRQKV NENPLYAITS EHVNVLLTTR EDRLVWKTQ GDPWKQLTN VPKGNFOATI EDWAEHKKGI
1001   IAAINSAPR TNPFSCKTNV CWAQRLEPIL ATAGIVLTGC QWSELPQFA DDKPHSAIYA LDVICIFFG MDLTSGLFSK QSIPLTYHPA DSARPVAHWD
1101   NSPGTRKYGY DHAVAAELSR RFPVFQLAGK GTQLDLQTGR TRVISAQHNL VPMNRNLPHA LVPEHKEKQP GPVKKFLSQF KHHSVLVSE EKIEAPHKRI
1201   EWIAPIGAG ADKNYNLAFG FPPQARYDLV FINIGTKYRN HHFQOCEDHA ATLKTLRSA LNCLNPGGTL VVKSYGYADR NSEDVVTALA RKFVRVSAAR
1301   PECVSSNTEM YLIFRQLDNS RTRQFTPHHL NCVISSYVEG TRDVGGAAPS YRTKRENIAD CQEEAVVNAA NPLGRPGEGV CRAIYKRWPN SFTDSATETG
1401   TAKLTVCGKG KVIHVGPDF RKHPAEALK LLQAYHAVA DLVNEHNKS VAIPLLSTGI YAAGKDRLEV SLNCLTLALD RTDADVTIYC LDKKWKERID
1501   AVLQKESVI ELKDEDMEID DELVWIHPDS CLKGRKGFTS TKGKLYSYFE GTKFHQAAD MAEIKVLFPN DQESNEQLCA YILGETMEAI REKCPVDHNP
1601   SSSPKTLPC LCMYAMTPER VHRLRSNNVK EVTVCSSTPL PKYKIKNVQK VQCTKVVLFN PHTPAFVPAR KYIEAPEQA APPAQAEAP EVAATPTPPA
1701   ADNTSLDVTD ISLDMESSE GSLFSSFGS DNSITSMDSW SSGPSSLEIV DRRQVVVADV HAVQEPAPVP PRLKKMARL AAARMQEEPT PPASTSSADE
1801   SLHLSFGGVS MSFGLFDGE MGALAAAQPP ASTCPTDVPD SFGSFDGEI EELSRRVTES EPVLFSGFEP GEVNSIISR SVVSFPPRKQ RRRRRSRRT
1901   Y

```

B. Amino Acid Sequence of the Structural Polyprotein

```

1      MNRGFFNMLG RRPFPAPTAM WRPRRRRQAA PMPARNGLAS QIQQLTTAVS ALVIGQATRP QTPRPRPPR QKKQAPKQPP KPKPKTQEK KKKQPAKPKP
101    GKRQRMALKL EADRLFDVKN EDGDVIGHAL AMEGKVMKPL HVKGTIDHPV LSKLKFTKSS AYDMEFAQLP VNMRSFAFTY TSEHPGFGYN WHHGAVQYSQ
201    GRFTIPRGVG GRGDSGRPIM DNSGRVVAIV LGGADEGTRT ALSVVTWNSK GKTIKTTPEG TEEWSAAPLV TAMCLLGNVS FPCNRPPTCY TREPSRALDI
301    LEENVNHEAY DTLNLAILRC GSSGRSKRSV TODFTLTPY LGTCSYCHHT EPCFSPIKIE QVWDEADDNT IRIQTSAQFG YDQSGAASN KYRYSLEQD
401    HTVKEGTMDD IKISTSGPCR RLSYKGYFLL AKCPPGDSVT VSIASSNSAT SCTMARKIKP KFGREKYDL PPVHGKKIPCV TVYDRLEKET AGYTTHMRPG
501    PHAYTSYLEE SSGKVYAKPP SGKNTIYECK CGDYKTGTVT TRTEITGCTA IKQCVAYKSD QTKWVFNPD LIRHADHTAQ GKLHLPFKLI PSTCMVPPAH
601    APNVVHGFKH ISLQDLDHLL TLLTTRRLGA NPEPTTEWII GKTVRNFTVD RDGLEIYWGN HEPVRVYAE SAPGDPHGWP HEIVQHYHYR HPVYTILAVA
701    SAAVAMMIGV TVAALCACKA RRECLTPYAL APNAVITSL ALLCCVRSAN AETFTETMSY LWSNSQPFVW VQLCIPLAAV IVLMRCCSCC LPFLVVAGAY
801    LAKVDAYEHA TTVPNVPQIP YKALVERAGY APLNLEITVM SSEVLPTNG EYITCKFTTV VSPKVKCCG SLEQCPAAHA DYTCVKFGGV YPFMWGGAQC
901    FCDSENSQMS EAYVELSADC ATDHAQAIV HTAAMKVGRL IVYGNTTSFL DVYVNGVTPG TSKDLKVIAG PISASFTPF HKVVIHRGLV YNYDFPEYGA
1001   MKPGAFGDIQ ATSLTSKDLI ASTDIRLLK SAKNVHVPYT QAASGFEMWK NNSGRPLQET APFGCKIAVN PLRAVDCSYG NIPISIDIPN AAFIRTSADP
1101   LVSTVKCDVS ECTYSADFGG MATLQYVSDR EGQCPVHSHS STATLQESTV HVLEKGAVTV HFSTASPAQAN FIVSLCGKKT TCNAECKPPA DHIVSTPHKN
1201   DQEFQAAISK TSWSWLFALF GGASSLLIG LMIFACSMML TSTRR

```

Fig. 4

Nucleotide Sequence of S55

1 ATTGGCGGCG TAGTACACAC TATTGAATCA AACAGCCGAC CAATTGCACT ACCATCACAA TGGAGAAGCC AGTAGTTAAC GTAGACGTAG ACCCTCAGAG TCCGTTTGTG GTGCAACTGC
 121 AAAAGAGCTT CCCGCAATTT GAGGTAGTAG CACAGCAGGT CACTCCAAAT GACCATTGTA ATGCCAGAGC ATTTTCGCAT CTGGCCAGTA AACTGTATGA GCTGGAGGTT CCTACCCAGC
 241 CGACGATTTT GGACATAGGC AGCGACCCGG CTCTAGAAAT GTTTTCCGAG CACCACTACC ATTGCGTTTG CCCCATGCGT AGTCCAGAAG ACCCGGACCG CATGATGAAA TATGCCAGCA
 361 AACTGGCGGA AAAAGCATGT AAGATTACAA ACAAGAACTT GCATGAGAAG ATCAAGGAGC TCCGGACCGT ACTTGATACA CCGGATGCTG AAACGCCATC ACTCTGCTTC CACAACGATG
 481 TTACCTGCAA CACGCGTGCC GAGTACTCCG TCATGCAGGA CGTGATACAT AACGCTCCCG GAACATATTA CCACCAGGCT ATGAAAGGCC TCGGACCCCT GTACTGGATT GGCTTCGACA
 601 CCACCCAGTT CATGTTCTCG GCTATGGCAG GTTCTGACCC TGCATACAA ACCAACTGGG CCGACGAAAA AGTCTTTGAA GCGCGTAACA TCGGACTCTG CAGCACAAAG CTGAGTGAAG
 721 GCAGGACAGG AAAGTTGTG ATAATGAGGA AGAAGGAGTT GAAGCCCGGG TCACGGGTTT ATTTCTCGT TGGATGACA CTTTACCAG AACACAGAGC CAGCTTGAG AGCTGGCAGT
 841 TTCCATCGGT GTTCCACTTG AAAGGAAAGC AGTGTACAC TTGCGCTGT GATACAGTGG TGAGCTGCGA AGGCTACGTA GTGAAGAAAA TCACCATCAG TCCCGGGATE ACGGGAGAAA
 961 CCGTGGGATA CCGGTTTACA AACAAATAGG AGGGCTTTT GCTATGAAA GTTACCGATA CAGTAAAGG AGAAGGGTA TCGTTCCCGG TGTGACGTA TATCCCGGCC ACCATATCGG
 1081 ATCAGATGAC CGGCATAATG GCCACGGATA TCTACCTGA CGATGCACAA AAATCTCTGG TTGGCTGCAA CCAGCGAATC GTCAATTAAC GTAAGACTAA CAGGAACACC AATACCATGC
 1201 AAAATTACCT TCTGCCAATC ATTGCACAA GGTTCAGCAA ATGGGCCAAG GAGCGCAAGG AAGATCTTGA CAATGAAAAA ATGCTGGCCA CCAGAGAGCG CAAGCTTACA TATGGCTGCT
 1321 TGTGGCGGTT TCGCACTAAG AAATGCACT CGTTCTATCG CCCACCTGGA ACGCAGACCA TCGTAAAGT CCCAGCCTCT TTTAGCGCTT TCCCATGTC ATCCGTATGG ACTACCTCTT
 1441 TGCCCATGTC GCTGAGGCG AGATGAAAT TGGCATTACA ACCAAGAAG GAGGAAAAAC TGCTGCAAGT CCCGAGGAA TTAGTTATGG AGGCCAAGGC TGCTTTCGAG GATGCTCAGG
 1561 AGGAATCCAG AGCGGAGAA CTCCGAGAA CACTCCACC ATTATGGCA GACAAAGGTA TCGAGGCAGC TCCGGAAGTT GTCTGCGAAG TGGAGGGGCT CCAGGCGGAC ACCGGAGCAG
 1681 CACTCTGCGA AACCCCGCG GGTCAATGTA GGATAATACC TCAAGCAAT GACCGTATGA TCGGACAGTA TATCGTTGTC TCGCGCATCT CTGTGCTGAA GAACGTAAA CTGCGACCA
 1801 CACACCCGCT AGCAGACCAG GTTAAGATCA TAACGCACTC CGGAAGTCA GGAAGGTATG CAGTCGAACC ATACGACGCT AAAGTACTGA TCGCAGCAGG AAGTCCCGTA CCATGCGCAG
 1921 AATTCTTAGC ACTGAGTGAG AGCGCCACGC TTGTGTACAA CGAAGAGAG TTTGTGAACC GCAAGCTGTA CCATATTGCC ATGCACGGTC CCGCTAAGAA TACAGAAGAG GAGCAGTACA
 2041 AGGTTACAAA GCGCAGAGCT CGAGAAACAG AGTACGTGT TGACGTGGAC AAGAAGCGAT CGGTTAAGAA GGAAGAAGCC TCAGGACTTG TCTTTTCGGG AGAACTGACC AACCCCGCCT
 2161 ATCAGCAACT AGCTTTTGAG GGACTGAAGA CTGACCCCG GGTCCCGTAC AAGGTTGAAA CAATAGGAGT GATAGGCACA CCAGGATCGG GCAAGTCAGC TATCATCAAG TCAACTGTCA
 2281 CGGCACGTGA TCTGTATACC AGCGGAAAGA AAGAAAACTG CCGCAAAAT GAGCGGAGC TGCTACGGCT GAGGGGATG CAGATCACTG CGAAGCAGT GGATTGCGTT ATGCTCAAGC
 2401 GATGCCACAA AGCCGTAGAA GTCTGTATG TTGACGAAGC GTTCCGTTGC CACGAGGAG CACTACTTGC CTGATTGCA ATGCTCAGAC CCGCTAAGAA GGTAGTACTA TCGGAGAGCC
 2521 CTAAGCAATG CGGATTCTTC AACATGATGC AACTAAAGT ACATTTCAC CACCCTGAAA AAGACATATG TACCAAGACA TTCTACAAGT TTATCTCCCG ACGTTGCACA CAGCCAGTCA
 2641 CGGCTATTGT ATCGACACTG CATTACGATG GAAAAATGAA AACCAACAAC CCGTCAAGA AGAATCAGA AATCGACATT ACAGGGGCCA CGAAGCCGAA GCCAGGGGAC ATCATCTGTA
 2761 CATGTTTCGG CGGTGGGTT AAGCAACTGC AAATCGACTA TCCCGGACAT GAGGTATGA CAGCCCGGCC CTCACAAGG CTAAACGAAA AAGGAGTATA TGCGTCCCG CAAAAAGTCA
 2881 ATGAAAAACC CCGTACGCG ATCACATCAG AGCATGTGAA CGTGTGCTC ACCCGCACTG AGGACAGGCT AGTATGAAA ACTTTACAGG CGGACCCATG GATTAAAGCAG CTCACTAAGC
 3001 TACCTAAAGG AAATTTTCAG GCCACCATCG AGGACTGGGA AGCTGAACAC AAGGGAATAA TTCTGCGAT AAACAGTCCC GCTCCCGTA CCAATCCGT CAGCTGCAAG ACTAACGTTT
 3121 GCTGGGGCAA AGCATGGAA CCGATACGG CCACGGCCGG TATGTAATT ACCGGTTGCC AGTGAGGCGA GCTTTTCCA CAGTTTCCGG ATGACAAACC ACATCGGCC ATCTACGCT
 3241 TAGAGTAAT TTGCATTAG TTTTCCGCA TGGACTGAC AAGCGGGCTG TTTTCCAAAC AGAGCATCCC GTTAACGTAC CATCTGCCG ACTCAGCGAG GCCAGTAGCT CATTGGGACA
 3361 ACAGCCGAGG AACACGCAAG TATGGTAGC ATCAGCGCGT TCGCGCGGAA CTCTCCGTA GATTTCGGT GTTCCAGCTA GCTGGGAAAG GCACACAGT TGAATTGCG AGCGGCAGAA
 3481 CTAGAGTTAT CTCTGCACAG CATAACTTGG TCCAGTGAA CCGCAATCTC CCTCAGGCT TAGTCCCGCA GCACAAGGAG AAACAACCCG GCGCGGTGGA AAATTTCTG AGCCAGTTCA
 3601 AACACCCTC CGTACTTGT ATCTCAGAGA AAAAAATGA AGTCCCCAC AAGAGAATCG AATGGATCG CCCGATTGGC ATAGCCGGCG CAGATAAGAA CTACAACCTG GCTTTCGGT
 3721 TTCCCGCGCA GGCACGGTAC GACCTGGTGT TCATCAATAT TGGAACTAAA TACAGAAACC ATCACTTCA ACAGTCCGAA GACCACCGCG CGACCTTGA AACCCCTTCC GCTTGGCGCC
 3841 TGAAGTCCCT TAACCCCGGA GGGACCCCTG TGGTGAATG CTACGGTTAC GCGGACCGCA ATAGTGAGGA CGTAGTACC GCTTTTCCA GAAAAATGTT CAGAGTGCTC GCAGCGAGG
 3961 CAGAGTCCGT CTCAGCAAT ACAGAAATG ACCTGATTT CCGACAATA GACAACAGCC GCACAGACA ATTCACCCCG CATCATTTGA ATTGTGTAT TCGTCCGTT TCGGAGGTA
 4081 CAAGAGACGG AGTTGAGGCC GCACCGCTGT ACCGTACTAA AAGGAGAAAC ATTGCTGATT GTCAAGAGGA AGCAGTTGTC AATGCAAGCA ATCCACTGGG CAGACAGGA GAAGGAGTCT
 4201 GCGTGGCAT CTATAAACGT TGGCCGAACA GTTTCACCGA TTCAGCCACA GAGACAGGTA CCGCAAACT GACTGTGTC CAAGGAAAGA AAGTATCCA CCGGGTTGGC CCGTATTTC
 4321 GGAACACACC AGAGGCAGAA GCCTGAAAT TGCTGCAAAA CGCTACCAT CGAGTGGCAG ACTTAGTAAA TGAACATAAT ATCAAGTCTG TCGCCATCCC ACTGCTATCT ACAGGCATT
 4441 ACGCAGCCGG AAAAGACCG CTGAGGTAT CACTTAACTG CTTGACAACC CGCTAGACA GAACTGATC GGAGTAAACC ATCTACTGCC TGGATAAGAA GTGGAAGGAA AGAATGAGC
 4561 CGGTGCTCCA ACTTAAGGAG TGTGTAAGT AGCTGAAGGA TGAGGATATG GAGATGAGC ACGAGTTAGT ATGGATCCAT CCGGACAGTT GCGTGAAGGG AAGAAAGGGA TTCAGTACTA
 4681 CAAAAAGGAAA GTTGTATTG TACTTTGAAG GCACCAAAAT CCATCAAGCA GCAAAAAGATA TGGCGGAGAT AAAGTCTCTG TTCCCAATG ACCAGGAAAG CAACGAACAA CTGTGTGCT
 4801 ACATATTGGG GGAGACCATG GAAGCAATCC GCGAAAAATG CCGGTGTCAC CACAACCCGT CGTETAGCCC GCCAAAAAGC CTGCGTCCG TGTGTATGTA TGCCATGAGC CCAGAAAGGG
 4921 TCCACAGACT CAGAAGCAAT AAGCTCAAG AAGTTACAGT ATGCTCTCC ACCCCCTTCC CAAAGTACA AATCAAGAA GTTCAGAAAG TTGAGTCCAC AAAAGTAGTC CTGTTTAAAC
 5041 CGCATACCCC CGCATTCGTT CCGCCCGTA AGTACATAGA AGCACCAGAA CAGCTGCGAG CTCCGCTGC ACAGGCGGAG GAGGCCCCCG GAGTTGAGC GACACCAACA CCACCTGCGC
 5161 CTGATAACAC CTGCTTGAT GTACCGGACA TCTCACTGGA CATGGAAGC AGTAGGCAAG GCTCACTTT TCGAGCTTT AGCGGATCCG ACAACTACCG AAGGCAAGTG GTGGTGGCTG
 5281 ACGTCCATGC CGTCCAAGAG CTTGCCCTG TTCCACCGCC AAGGCTAAG AAGATGCCG GCTGGCAGC GGCAAGAATG CAGGAAGAGC CAATCCACC GGCAGGACC AGCTCTCGCG
 5401 ACGAGTCCCT TCACCTTTCT TTTGATGGG TATCTATATC CTTCGGATCC CTTTTCGAGC GAGAGATGGC CCGCTTGGCA GCGGCACAA CCCCCGCAAG TACATGCCCT ACGGATGTC
 5521 CTATGCTTTT CGGATCGTT TCCGACGGAG AGATTGAGGA GTTGAGCCCG AGAGTAAACC AGTGGAGCC GCTCTGTTT GGTCTATTG AACCGGGCGA AGTGAATCA ATTATATCT
 5641 CCGCATCAGC CGTATCTTT CCACCACGA AGCAGAGAGC TAGACGAGG AGCAGGAGGA CCGAATACGT TCTAACCGGG GTAGGTGGGT ACATATTTTC GACGGACACA GCGCCGTGGC
 5761 ACTTGCAAAA GAAGTCCGT CTGCGAAGC AGCTTACAGA ACCGACCTTG GAGCGCAATG TTCTGGAAG AATCTACGCC CCGGTGCTCG ACACGTGCAA AGAGGAACAG CTCAAACTCA
 5881 GGTACCAGAT GATGCCCAAC GAAGCAACA AAAGCAGGTA CCAATCTCGA AAAGTAGAAA ACCAGAAAGC CATAACCACT GAGCGACTGC TTTCAGGGCT ACGGCTGTAT AACTTGCCA
 6001 CAGATCAGCC AGAATGCTAT AAGTACCTT ACCCGAAACC ATGATTTCC AGCAGTGTAC CAGCGAACA CTGTGACCCA AAGTTTCTG TAGCTGTTT TAACAACATAT CTGATGAGA
 6121 ATTACCCGAC GGTAGCATCT TATCAGATCA CCGACGAGTA CGATGCTTAC TTGGATATGG TAGACGGAC AGTCCCTTGC CTAGATACT CAACCTTTTG CCGCGCCAG CTTAGAGTT
 6241 ACCCGAAAAG ACAGAGTAT AGACCCCAA ACATCCCGAG TCGGTTTCCA TCAGGATGC AGAACAGTT GCAAAAGCTG CTCATTGCGC CGACTTAAAG AAATGCAAC GTACACAAAA
 6361 TCGTGAACT GCAACACTG GACTCAGGCA CATTAACGT TGAATGCTT CGAAAATATG CATGCAATGA CGAGTATTG GAGGAGTTG CCGGAAAGCC AATTAGGATC ACTACTGAT
 6481 TCGTTACCGC ATACGTGGCC AGACTGAAA GCGCTAAGC GCGCGCACTG TTGCAAGA CCGATAATTT GGTCCCATG CAAGAAGTGC CTATGGATAG ATTCGTCATG GACATGAAA
 6601 GAGACGTGAA AGTTACACTT GGCACGAAAC ACACAGAAGA AAGACCGAAA GTACAAAGTA TACAAGCCGC AGAACCCCTG GCGACCGCTT ACCTATGCGG GATCCACCGG GAGTAGTGC

Fig 5A

6721 GCAGGCTTAC AGCGTTTGT CTACCAACA TTCACCGCT CTTTGACATG TCGCGGAGG ACTTTGATGC AATCATAGCA GAACACTTCA AGCAAGGTGA CCCGGTACTG GAGACGGATA
6841 TCGCTCTGTT CGACAAAAGC CAAGACGACG CTATGCGTT AACCGGCTG ATGATCTTGG AAGACCTGGG TGTGGACCAA CCACTACTCG ACTTGATCGA GTGCGCTTT GGAGAAATAT
6961 CATCEACCCA TCTGCCACG GGTACCGTT TCAAAATCGG GCGGATGAT AAATCCGAA GTTCTCTAC GCTCTTTGTC AACACAGTTC TGAATGTCTG TATCGCCAGC AGAGTATGG
7081 AGGAGCGGCT TAAACGTCC AAATGTGCG CATTTATCGG CGACGACAAC ATTATACACG GAGTAGTATC TGACAAAGAA ATGGCTGAGA GGTGTGCCAC CTGGCTCAAC ATGGAGGTGA
7201 AGATCAATTGA CGCAGTCATC GCGGAGAGAC CACCTTACTT CTGCGGTGGA TTCTATTGCG AAGATTCGGT TACCTCCACA GCGTGTGCGG TGGCGGACCC CTGAAAAAGG CTGTTAAAGT
7321 TGGGTAAACC GCTCCACGCC GACGATGAGC AAGACGAAGA CAGAAGACGC GCTCTGCTAG ATGAACAAA GCGCTGTTT AGAGTAGGTA TAACAGACAC CTTAGCAAGTG GCGGTGGCAA
7441 CTCGGTATGA GGTAGACAAC ATCACACCTG TCTGCTGGC ATTGAGAACT TTTGCCCAGA GCAAAAGAGC ATTCAAGCC ATCAGAGGGG AAATAAGCA TCTCTACGGT GGTCTAAAT
7561 AGTCAGCATA GTACATTCA TGTGACTAAT ACCACAACAC CACCACGATG AATAGAGGAT TCTTTAATC GCTCGGCCG CCGCCCTTCC CAGCCCCCAG TGCATGTGG AGGCGCGGGA
7681 GAAGGAGGCA GCGCGCCCGC ATGCTGCCG CAAATGGCT GGTCTCCAA ATCCAGCAAC TGACCACAGC CGTCAGTGGC CTAGTCAATG GACAGGCAAC TAGACCTCAA ACCCCACGCC
7801 CACGCCCGCC GCGCGCCGAG AAGAAGCAGG CGCCAAAGCA ACCACCGAAG CCGAAGAAAC CAAAAACACA GGAGAAGAAG AAGAAGCAAC CTGCAAAACC CAAACCCGGA AAGAGACAGC
7921 GTATGGCACT TAAATGGAG GCGGACAGAC TGTTCGAGCT CAAAAATGAG GACGGAGATG TCATCGGGCA CGCACTGGCC ATGGAAGGAA AGGTAATGAA ACCACTCCAC GTGAAGGAA
8041 CTATTGACCA CCCTGTGCTA TCAAAAGCTA AATTEACCAA GTGCTACGCA TACGACATGG AGTTCGACA GTTCCCGTC AACATGAGAA GTGAGGCGTT CACCTACCC AGTGAACACC
8161 CTGAAGGGTT CTACAACCTG CACCACGGAG CGGTGCACTA TAGTGGAGGC AGATTACCA TCCCCCGCG AGTAGGAGGC AGAGGAGACA GTGCTGCTCC GATTATGGAT AACTACGGCC
8281 GGGTTGTGCG GATAGTCTC GGAGGGGCTG ATGAGGGAAC AAGAACCACC CTTCGGTGG TCACCTGGA TAGCAAGGG AAGACAATCA AGACAACCC GGAAGGGACA GAAGAGTGT
8401 CTGCTGACCC ACTGTGACG GGCATGTGCT TGTGTGAAA CGTGAGCTTC CCATGCAATC GCGCGCCAC ATGCTACACC CCGGAACCAT CCAGAGCTTT CGACACTCTC GAAGAGAAGC
8521 TGAACCACGA GCGCTACGAG ACCCTGCTCA ACCGCATATT CCGGTGCGGA TGTGCGGCA GAAGTAAAG AAGCGTCACT GACGACTTTA CTTTGACCAG CCGGTCTTG GGCACATGCT
8641 CGTACTGTCA CCATACTGAA CCGTCTTTA GCGCGATTAA GATCGAGCAG GTCTGGGATG AAGCGGACGA CAACACCATA CGCATACAGA CTTCGCCCCA GTTTGGATAC GACCAAGCG
8761 GAGCAGCAAG CTCAAATAAG TACCGCTACA TGTGCTCGA GCAGGATCAT ACTGTCAAAG AAGGCACCAT GGATGACATC AAGATCAGCA CCTCAGGACC GTGTAGAAGG CTTAGCTACA
8881 AAGGATACTT TCTCTCGCG AAGTGTCTC CAGGGGACAG CGTAACGGTT AGCATAGGGA GTAGCAACTC AGCAACGTCA TGCACAATGG CCGCAAGAT AAAACCAAAA TGTGTGGAC
9001 GGGAAAAATA TGACCTACCT CCGGTTCAGG GTAAGAAGAT TCTTGCACA GTGTACGACC GTTGAAGA AACAACCGCC GGTACATCA CTATGCACAG GCGGGGACCG CACGCTATA
9121 CATCTATCT GGAGGAATCA TGAGGGAAG TTTACGGGA GCCACCATCC GGAAGAACA TTACGTACGA GTGCAAGTGC GCGGATTACA AGACCGGAAC CGTTACGACC CGTACCGAAA
9241 TCACGGGCTG CACCGCCATC AAGCAGTGGC TCGCTATAA GAGCGACCAA ACGAAGTGGG TCTTCAACTC GCGGACTCG ATCAGACAGC CCGACCAAC GCGCAAGGG AAATTCGATT
9361 TGCCTTTCAA GCTGATCCG AGTACCTGCA TGTCTCTGT TCGCCACGGC CGGAACGTAG TACACGGCTT TAAACACATC AGCTTCAA TAGACACAGA CCATCTGACA TTGCTEACCA
9481 CCAGGAGACT AGGGGCAAC CCGGAACCAA CCACTGAATG GATCATGGA AACACGGTGA GAACTTCACT CGTCGACCGA GATGGCTGG AATACATATG GGGCAATCAC GAACAGTAA
9601 GGTCTATGC CCAAGATCT CCACCAAGAG ACCCTEACGG ATGGCCACAC GAAATAGTAC AGCATTAATA TCATCGCCAT CCTGTGTACA CCATCTTAGC CGTCGCATCA GCTGTGTG
9721 CGATGATGAT TGGCGTAACT GTTCAGCAT TATGTGCTG TAAAGCGCG CGTAGTGGC TGACGCCATA TCGCTGGCC CCAATGGCG TGATTECAAC TTCGTGGCA CTTTGTGCT
9841 GTGTAGGTC GGTAAATGCT GAAACATCA CCGAGACCAT GAGTTACTTA TGTGGAACA GCCAGCGTT CTCTGGGTC CAGCTGTGTA TACCTGTGCG CGCTGTGCT GTTCTAATGC
9961 GCTGTGCTC ATGCTGCGTG CTTTTTTAG TGGTTGCGG CGCTACCTG CGGAAGTAG ACGCTACGA ACATGGGACC ACTGTTCGA ATGTGCCACA GATACCGTAT AAGGCACCTG
10081 TTGAAGGGC AGGTACGCC CCGCTCAATT TGGAGATTAC TGTATGTCC TGGAGGTTT TCGCTTCCAC CAACCAAGAG TACATTACCT GCAAAATCACT CACTGTGCTC CCCTCCCTA
10201 AAGTCAGATG CTGCGGCTCC TTGGAATGTC AGCCCGCCCG TCACGACAGC TATACCTGCA AGGTCTTTGG AGGGGTGTAC CCCTTCATGT GGGGAGGAGC ACAATGTTTT TCGCAGATG
10321 AGAACAGCCA GATGAGTGA GGTACGTGCT AATTGTCACT AGATTGCGCG ACTGACCAGC CGCAGGCGAT TAAGGTGCACT ACTGCGCGCA TGAAGTAGG ACTGCGTATA GTGTACGGGA
10441 ACATACCCAG TTTCTAGAT GTGTACGTA ACGGATGAC ACCAGGAACG TCTAAGAGCC TGAAGTCAAT AGCTGGACCA ATTTACGAT TGTTTACACC ATTCGATCAC AAGGTGTTA
10561 TCAATCGCGG CCGGTGTAC AACTATGACT TTCCGGAATA CGGAGCGATG AAACAGGAG CGTTTGGAGA CATTCAAGCT ACCTCTTTGA CTAGCAAGA CTTCTCGCC AGCAGAGACA
10681 TTAGGCTACT CAAGCTTCC GCAAGAAGC TGCAATGCTC GTACAGCGAG GCGCATCTG GATTCGAGAT GTGAAAAAC AACTCAGGCC GCGCATGCA GGAACCGCC CTTTTGGGT
10801 GCAAGATTGC AGTCAATCCG CTTCGAGCGG TGGACTGCTC ATACGGGAAC ATTCCTATT TATTGACAT CCGGAACGCT GCCTTTATCA GGACATCAGA TGACCACTG GTCTCAACAG
10921 TCAAAATGTA TGTCACTGAG TGCACTTATT CAGCGGACTT CGGAGGATG GCTACCTGC AGTATGTAT CGACCGCGAA GGACAATGCC CTGTACATTC GCATTCGAGC ACAGCAACCC
11041 TCCAAGAGTC GACAGTTTAT GTCTGGAGA AAGGAGCGGT GACAGTACAC TTCAGCACCG CGAGCCACA GCGGAACCTC ATTGTATGCG TGTGTGTAA GAAGACAACA TCCAATGCA
11161 AATGCAAAAC ACCAGTGTAT CATATCTGTA GCACCCCGCA CAAAAATGAC CAAGAATTC AACCGCCAT CTCAAAACT TCATGGAGTT GGTGTGTC CTTTTGCGG GCGGCTCTG
11281 CGCTATTAA TATAGGACTT ATGATTTTT CTTCAGCAT GATGCTGACT AGCAGACGAA GATGACCGCT ACGCCCAAT GACCCGACCA GCAAAACTG ATGTACTTCC GAGGAACCTA
11401 TGTGATAAT GCATCAGGCT GGTATATTAG ATCCCGCTT ACCCGGGCA ATATAGCAAC ACCAAAACTC GACGTATTTC CGAGGAAGCG CAGTGCAATA TGTGCGCAG TGTGCGCAA
11521 TAATCACTAT ATTAACCAT TATTCAGCG AGCCAAAAAC TCAATGATT TGTGAGGAG CATGTGCAAT AATGCCATGC AGCGTCTGCA TAACTTTTA TTATTTCTT TATTAATCAA
11641 CAAAATTTT TTTTAAATC TTC

Fig. 5 B

Nucleotide Sequence of TR339

1 ATTGGCGGCG TAGTACACAC TATTGAATCA AACAGCCGAC CAATTGCACT ACCATACAA TGGAGAAGCC AGTAGTAAAC GTAGACGTAG ACCCCAGAG TCCGTTTGTG GTGCAACTGC
121 AAAAAAGCTT CCCGCAATTT GAGGTAGTAG CACAGCAGGT CACTCCAAT GACCATGCTA ATGCCAGAGC ATTTTCGCAT CTGGCCAGTA AACTAATGGA GCTGGAGGTT CCTACCACAG
241 CGACGATCTT GGACATAGGC AGCGCACCAG CTCGTAGAAT GTTTCCGAG CACCAGTATC ATTGTGCTG CCCCATGCGT AGTCCAGAAG ACCCGGACCG CATGATGAAA TATGCCAGTA
361 AACTGGCGGA AAAAGCGTGC AAGATTACAA ACAAGAACTT GCATGAGAAG ATTAAGGATC TCCGACCGT ACTTGATACG CCGGATGCTG AAACACCATC GCTCTGCTTT CACAACGATG
481 TTACCTGCAA CATCGTGC CC GAATATCCG TCATGCAGGA CGTGATATC AACGCTCCCG GAACTATCTA TCATCAGGCT ATGAAAGGCG TCGCGACCCT GTACTGGATT GGTCTGCACA
601 CCACCAGATT CATGTTCTCG GTATGGCAG GTTCGTACCC TCGGTACAA ACCAACTGGG CCGACGAGAA AGTCCTTGA CCGCGTAACA TCGGACTTTG CAGCACAAAG CTGAGTGAAG
721 GTAGGACAGG AAAATTGTGC ATAATGAGGA AGAAGGAATT GAAGCCCGGG TCGCGGTTT ATTTCTCCGT AGGATCGACA CTTTATCCAG AACACAGAGC CAGCTTGACG AGCTGGCAGC
841 TTCCATCGGT GTTCCACTTG AATGAAAGC AGTCGTACAC TTGCGCGTGT GATACAGTGG TGAGTTGCGA AGGCTACGTA GTGAAGAAAA TCACCATCAG TCCCGGGATC ACGGAGAGAA
961 CCGTGGGATA CCGGTTTACA CACAATAGCG AGGGCTCTTT GCTATGCAAA GTTACTGACA CAGTAAAGG AGAACGGGTA TCGTCCCTG TGTGACGTA CATCCCGGCC ACCATATGCG
1081 ATCAGATGAC TGTATAATG GCCACGGATA TATCACTGA CGATGCACAA AAATCTCTCG TTGGGCTCAA CCAGCGAATT GTCTTAACG GTAGGACTAA CAGGAACACC AACACCATGC
1201 AAAATTACCT TCTCCGAGAA ATAGCAAAAG GGTTCAGCAA ATGGGCTAAG GAGCGCAAGG ATGATCTTGA TAACGAGAAA ATGCTGGGTA CTAGAGAAGC CAAGCTTACG TATGGCTGCT
1321 TGTGGCGT TTGCACTAAG AAATGATCAT CTTTTATCG CCCACCTGGA ACGCAGACCA TCGTAAAAAT CCCAGCCTCT TTAGCGCTT TTCCATGTC TCGGTATGG ACGACCTCTT
1441 TCGCCATGTC GCTGAGGCG AAATGAAAC TGGCATTGCA ACCAAAGAG GAGGAAAAAC TGCTGCAGT CTEGAGGAA TTAGTCATGG AGGCCAAGGC TGCTTTTGG GATGCTCAGG
1561 AGGAAGCCAG AGCGGAGAG CTCCGAGAAG CACTTCCACC ATTAGTGGCA GACAAAGGCA TCGAGGCAGC CGCAGAAGTT GTCTGCGAAG TGGAGGGGCT CCAGGCGGAC ATCGGAGCAG
1681 CATTAGTTGA AACCCCGCG GGTACGTAA GGATAATACC TAAAGCAAT GACCGTATGA TCGGACAGTA TATGTTGTC TCGCCAACT CTGTCTGAA GAATGCCAAA CTCGCCAG
1801 CGCACCCGCT AGCAGATCAG GTTAAGATCA TAACACACTC CGGTAGATCA GGAAGGTACG CGGTGCAACC ATACAGCCT AAAGTACTGA TGCCAGCAGG AGGTGCGCTA CCATGGCCAG
1921 AATTCTAGC ACTGAGTGAG AGCGCCACGT TAGGTACAA CGAAAGAGAG TTTGTGAACC GCAAACTATA CCACATTGCC ATGATGGCC CCGCCAAAG TACAGAAGAG GAGCAGTACA
2041 AGGTTACAAA GGCAGAGCTT GCAGAAACAG AGTACGTGTT TGACGTGGAC AAGAAGCGTT GCGTTAAGAA GGAAGAAGCC TCAGGTCTGG TCTCTCGG AGAACTGACC AACCTCCCT
2161 ATCATGAGCT AGCTCTGGAG GGACTGAAGA CCCGACCTGC GGTCCGTAC AAGGTGAAA CAATAGGAGT GATAGGCACA CCGGGGTGCG GCAAGTCAGC TATTATCAAG TCAACTGTCA
2281 CGGCACGGGA TCTTTTACC AGCGGAAAG AAGAAATTG TCGGAAATT GAGGCGGAG TGCTAAGACT GAGGGGTATG CAGATTACGT CGAAGACAGT AGATTCCGTT ATGCTCAACG
2401 GATGCCACAA AGCCGTAGAA GTTCTGTACG TTGACGAAGC GTTCCGTGCG CACGACGAGG CACTACTTGC CTGTAGTCT ATGCTCAGGC CCGCCAAAG GGTAGTACTA TCGGAGAGAC
2521 CCATGCAATG CGGATTCTTC AACATGATGC AACTAAAGT ACATTTCAT CACCCTGAAA AAGACATATG CACCAAGACA TTCTACAAGT ATATCTCCCG GCGTTGACA CAGCCAGTTA
2641 CAGCTATTGT ATCGACACTG CATTACGATG GAAAGATGAA AACCCAGAAC CCGTGCAAGA AGAACATTGA AATCGATATT ACAGGGGCGA CAAGCCGGA GCCAGGGGAT ATCATCTGA
2761 CATGTTTCCG CCGGTGGGTT AAGCAATTGC AAATCGACTA TCCCGGACAT GAAGTAATGA CAGCCGCGCG CTCACAAGGG CTAACCAGAA AAGGAGTGTG TGCCGTCCGG CAAAAAGTCA
2881 ATGAAAAACC ACTGTACGCG ATCAGATCAG AGCATGTGAA CGTGTGCTC ACCCGCACTG AGGACAGGCT AGTGTGAAA ACCTTGACAG GCGACCCATG GATTAGCAG CTCACTAACA
3001 TACCTAAAGG AAATTTTCA GCTACTATAG AGGACTGGGA AGCTGAACAC AAGGGAATAA TTGCTGCAAT AACAGCCCC ACTCCCGTG CCAATCCGTT CAGCTGCAAG ACCAAGCTTT
3121 GCTGGGCGAA AGCATTGGA CCGATACTAG CCACGGCCGG TATCGTACTT ACCGGTGC AGTGGAGCGA ACTGTTCCA CAGTTTGGG ATGACAAACC ACATTCGGCC ATTTACGCT
3241 TAGACGTAAT TTGCATTAAG TTTTCCGCA TGGACTTGAC AAGCGGACTG TTTTCTAAAC AGAGCATCCC ACTAACGTAC CATCCCGCG ATTACGCGAG GCCGGTAGCT CATTGGGACA
3361 ACAGCCGAGG AACCCGCAAG TATGGGTACG ATCAGCCCAT TCCCGCGGAA CTCTCCCGTA GATTTCGGT GTTCCAGCTA GCTGGGAAGG GCACACAAT TGAATTCGAG ACGGGGAGAA
3481 CCAGAGTTAT CTCTGCACAG CATAACCTGG TCCCGGTGAA CCGCAATCTT CACTACGCTT TAGTCCCGA GTACAAAGG AAGCAACCCG GCGCGGTGGA AAAATCTTGT AACCAATTCA
3601 AACACCACCT AGTACTTGTG GTATCAGAGG AAAAAATTGA AGTCCCGCT AAGAGATCG AATGGATCG CCGCATGGC ATAGCCGGTG CAGATAAGAA CTACAACCTG CTTTTCGGGT
3721 TCCCGCCGCA GGCACGGTAC GACCTGGTGT TCATCAACAT TGGAACTAAA TACAGAAACC ACCACTTTCA GCAGTGCAGG GACCATGCGG CGACCTTAAA AACCTTTGCG CGTTCGGCCC
3841 TGAATTGCTT TAACCCAGGA GGCACCCCTG TGTGAAAGT CTATGGTAC GCGGACCGCA ACAGTGAGGA CGTAGTCACC GCTTTGCCA GAAAGTTTGT CAGGGTGTCC GCACGAGAC
3961 CAGATTGTGT CTCAGCAAT ACAGAAATGT ACCTGATTTT CCGACAATA GACAACAGCC GTACACGGCA ATTCACCCCG CACCATCTGA ATTGCGTAT TTGCTCGTG TATGAGGTA
4081 CAAGAGATGG AGTTGGAGCC GCGCGTCTAT ACCGACCAAA AAGGAGAAAT ATTGCTGACT GTCAAGAGGA AGCAGTTGTC AACGCAGCCA ATCCGCTGG TAGACAGGC GAAGGAGTCT
4201 CCGCTGCCAT CTATAACGT TGGCCGACCA GTTTTACCGA TTCAGCCACG GAGACAGGCA CCGCAAGAA GACTGTGTGC CTAGGAAAGA AAGTGATCCA CCGCGTGGC CTGATTTCC
4321 GGAAGCACCC AGAAGCAGAA GCCTTGAAT TGTACAAAA CGCTACCAT GCAGTGGCAG ACTTAGTAAA TGAACATAAC ATCAAGTCTG TCGCCATTCC ACTGCTATCT ACAGGCATTT
4441 ACGCAGCCGG AAAAGACCCG CTTGAAGTAT CACTTAACTG CTTGACAACC GCGGTAGACA GAACTGACGC GGACGTAACC ATCTATTGCC TGGATAAGAA GTGGAAAGGA AGAATCGAGC
4561 CCGCACTCCA ACTTAAAGG TCTGTAAAG AGCTGAAGGA TGAAGATATG GAGATGAGC ATGAGTTAGT ATGATTCAT CCAGACAGTT GCTTGAAGGG AAGAAAGGGA TTCAGTACTA
4681 CAAAAGGAAA ATTGTATTG TACTTCGAAG GCACCAAAAT CCATCAAGCA GCAAAAGACA TGGCGGAGAT AAAGTCTGT TTCCTAATG ACCAGGAAAG TAATGAACAA CTGTGTGCT
4801 ACATATTGGG TGAGACCATG GAAGCAATCC GCGAAAAGTG CCGGTGCGAC CATAACCCGT COTCTAGCC GCGCAAAAGG TTGCGGTGCC TTTGATGTA TGCCATGACG CCAGAAAGGG
4921 TCCACAGACT TAGAAGCAAT AACGTCAAAG AAGTTACAGT ATGCTCTCC ACCCCCCCTT CTAAGCACA AATTAAGAAT GTTCAGAAGG TTCAGTGCAC GAAAGTAGTC CTGTTAATC
5041 CGCACACTCC CGCATTCGTT CCGCCCGTA AGTACATAGA AGTGCCAGAA CAGCCTACCG CTCCTCTGCG ACAGGCGGAG GAGGCCCCCG AAGTTGTAGC GACACCGTCA CCATCTACAG
5161 CTGATAACAC CTCGCTTGAT GTCACAGACA TCTCACTGGA TATGGATGAC AGTAGCGAAG GCTCACTTTT TTCAGCTTT AGCGGATCGG ACAACTCTAT TACTAGTATG GACAGTTGCT
5281 CGTCAGGACC TAGTTCTA GAGATAGTAG ACCGAAGGCA GGTGGTGGT GCTGACGTTT ATGCGCTCCA AGAGCCTGCC CTTATTCCAC CGCCAAAGGT AAAGAAGATG GCCCGCCTGG
5401 CAGCGGCAAG AAAAGAGGCC ACTCCACCGG CAAGCAATAG CTCTGAGTCC CTCACCTCT CTITTTGGTG GGTATCCATG TCCTCCGGAT CAATTTTCCA CGGAGAGAGC GCCCGCCAGG
5521 CAGCGGTACA ACCCTGGCA ACAGGCCCCA CGGATGTGCC TATGCTTTT GGATCGTTTT CCGACGGAGA GATTGATGAG CTGAGCCGCA GAGTAACTGA GTCCGAACCC GTCTGTGTTG
5641 GATCATTTGA ACCGGCGGAA GTGAATCAA TTATATCTG CCGATCAGCC GTATCTTTT CACTACGCAA GCAGAGAGCT AGAGCAGGA GCAGGAGGAC TGAATACTGA CTAACCGGGG
5761 TAGTGGGTA CATATTTTCG ACGGACACAG GCCCTGGGCA CTTGCAAAAG AAGTCCGTTT TGCAGAACCA GCTTACAGAA CCGACCTTGG AGCGCAATGT CTTGAAAAA ATTCTAGCCC
5881 CGGTGCTGCA CAGCTGAAA GAGGAACAAC TCAAACTCAG GTACCAGATG ATGCCACCG AAGCAACAA AAGTAGGTAC CAGTCTCGTA AAGTAGAAAA TCAGAAAGCC ATAACCACTG
6001 AGCGACTACT GTCAGGACTA CGACTGTATA ACTCTGCCAC AGATCAGCCA GAATGTATA AGATCAGCTA TCCGAAACCA TTGTACTCCA GTAGCGTACC GCGGAATAC TCGATCCAC
6121 AGTTCGCTGT AGCTGTCTGT AACAATATC TGCATGAGAA CTATCCGACA GTAGCATCTT ATCAGATTAC TGACGATGAC GATGCTTACT TGGATATGTT AGACGGGACA GTCCCTGCTC
6241 TGAATACTGC AACCTTCTGC CCGCTAAGC TTAGAAGTA CCGGAAAAAA CATGAGTATA GAGCCCCGAA TATCCGAGT CCGGTTCAT CAGCGATGCA GAACACGCTA CAAAATGTGC
6361 TCATTGCCCG AACTAAAAAG AATTGCAACG TCACGAGAT CCGTGAAGTG CCAACACTGG ACTCAGGCAC ATTCATGTCT GAATGCTTTC GAAAATATGC ATGTAATGAC GAGTATTGGG
6481 AGGAGTTCGG TCGGAAGCCA ATTAGGATTA CCACTGAGTT TGTACCCGCA TATGATGCTA GACTGAAAGG CCTAAGGCC GCGGCACTAT TTGCAAGAGC GTATAATTTG GTCCCATTCG
6601 AAGAAGTGCC TATGGATAGA TTGCTCATGG ACATGAAAG AGACGTGAAA GTTACACAGG GCACGAAACA CACAGAAGAA AGACCGAAG TACAAGTAT ACAAGCCGCA GAACCCCTGG

Fig 6A.

6721 CGACTGCTTA CTTATGCGGG ATTCACCGGG AATTAGTGG TAGGCTTACG GCCGTCTTGC TTCCAAACAT TCACACGCTT TTGACATGT CGGCGGAGGA TTTTGATGCA ATCATAGCAG
6841 AACACTTCAA GCAAGGCGAC CCGTACTGG AGACGGATAT CGCATEATTC GACAAAAGCC AAGACGACGC TATGGCGTTA ACCGGTCTGA TGATCTTGA GGACCTGGGT GTGGATEAAC
6961 CACTACTCGA CTTGATCGAG TCGGCTTTG GAGAAATATC ATCCACCCAT CTACCTACGG GTACTCGTTT TAAATTCGGG GCGATGATGA AATCCGGAAT GTTCTCTACA CTTTITGTCA
7081 ACACAGTTTT GAATGTCTTT ATGCCACGA GAGTACTAGA AGAGCGGCTT AAAACGTCCA GATGTGCAGC GTTCTATTGC GACGACAACA TCATACATGG AGTAGTATCT GACAAAAGAA
7201 TGGCTGAGAG GTGCCACCC TGGCTCAACA TGGAGTTAA GATCATCGAC GCAGTCACTG GTGAGAGACC ACCTTACTTC TCGGCGGAT TTATCTTGA AGATTEGGTT ACTTCCACAG
7321 CGTGCCGCTT GCGGACCCCT CTGAAAAGGC TGTTTAAGTT GGGTAAACCG CTCCCAGCCG ACGACGAGCA AGACGAAGAC AGAAGACGCG CTCTGTAGA TGAAACAAAG GCGTGGTTTA
7441 GAGTAGGTAT AACAGGCACT TTAGCAGTGG CCGTGACGAC CCGGTATGAG GTAGACAATA TTACACCTGT CCTACTGGCA TTGAGAACTT TTGCCAGAG CAAAAGAGCA TTCCAAGCCA
7561 TCAGAGGGGA AATAAGCAT CTCTACGGTG GTCTAAATA GTACGATAG TACATTTCAT CTGACTAATA CTACAACACC ACCACCATGA ATAGAGGATT CTTTAACATG CTCGGCCGCC
7681 GCGCTTCCC GCGCCCACT GCCATGTGA GCGCGCGGAG AAGGAGGAG GCGGCCCGGA TCGCTGCCCG CAACGGGCTG GCTTCTCAA TCCAGCACT GACCACAGCC GTCACTGCC
7801 TAGTCATTGG ACAGGCAACT AGACCTCAAC CCCACGTCC ACGCCGCCCA CCGGCCAGA AGAAGCAGGC GCCCAAGCAA CCACCGAAGC CGAAGAAACC AAAAACGAG GAGAAGAAGA
7921 AGAAGCAACC TGAAAAACC AAACCGGAA AGAGACAGCG CATGGCACTT AAGTTGGAGG CCGACAGATT GTTCGACGTC AAGAACGAGG ACCGAGATGT CATCGGGCAC GCACTGGCCA
8041 TGGAAAGAAA GGTAAATGAA CCTCTGCAGC TGAAAGGAAC CATEGACCAC CCGTGTCTAT CAAAGCTCAA ATTTACCAAG TCGTCAGCAT ACGACATGGA GTTCGACAG TTGCCAGTCA
8161 ACATGAGAAG TGAGGCATTC ACCTACACCA GTGAACACCC CGAAGGATTC TATAACTGGC ACCACGGAGC GGTGCAGTAT AGTGGAGGTA GATTTACCAT CCCTCGCGGA GTAGGAGGCA
8281 GAGGAGACAG CCGTCTGTCG ATCATGGATA ACTCCGTCG GTTGTGCGG ATAGCTCTCG GTGAGAGTGA TGAAGGAACA CGAATCGCCC TTTCGGTCT CACTGGAAT AGTAAAGGGA
8401 AGACAATTAA GACGACCCCG GAAGGACAG AAGAGTGTTC CCGACACCA CTGCTACCG CAATGTGTTT GCTCGGAAAT GTGAGCTTCC CATGCGACCG CCGGCCACA TGCTATACCC
8521 CGGAACCTTC CAGAGCCCTC GACATCTTG AAGAGAAGCT GAACCATGAG GCCTACGATA CCGTGTCAA TGCCATATTG CCGTGGGAT CCGTCTGCGA AAGCAAAAGA AGCGTCACTG
8641 ACGACTTAC CCGACACAG CCGTACTTGG GCACATGCTC GTACTGCCAC CATCTGAAC CCGTCTTAC CCGTGTAAAG ATGAGCAGG TCTGGGACGA AGCGGACGAT AACACCATC
8761 GCATACAGAC TTCCGCCAG TTTGGATAG ACCAAAGCGG AGCAGCAAGC GCAACAAGT ACCGTACAT GTGCTTGAG CAGGATCACA CCGTTAAAGA AGGCACCATG GATGACATCA
8881 AGATTAGCAC CTCAGGACCG TGTAGAAGC TTAGCTACAA AGGATACTTT CTCTCGCAA AATGCCCTEC AGGGACAGC GTAACGGTTA GCATAGTGA TAGCAACTCA GCAAGCTCAT
9001 GTACACTGGC CCGCAAGATA AAACCAAAAT TCGTGAGCG GGAATAATAT GATCTACCTC CCGTACCGG TAAAAAATTT CTTGACAGG TGTACGACCG TCTGAAAGAA ACAACTGCAG
9121 GCTACATCAC TATGACAGG CCGGACCGC ACGCTTATAC ATCCTACCTT AAGAATCAT CAGGGAAATG TTACGCAAAAG CCGCCATCTG GGAAGAATAT TACGTATGAG TGCAAGTGG
9241 GCGACTACAA GACCGGAACC GTTTCGACCC GCACGAAAT CACTGTTGC ACCGCCATCA AGCAGTGGCT CCGCTATAAG AGCGACCAAA CGAAGTGGT CTTCAACTCA CCGGACTTGA
9361 TCAGACATGA CGACCAACAG GCCCAAGGA AATTGCATTT GCCTTCAAAG TTGATCCGA GTACCTGCAT GGTCCTGTT GCCACGCGC CGAATGTAAT ACATGGCTTT AAACACATCA
9481 GCCTCAATT AGATACAGAC CACTTGACAT TGCTACCAC CAGGAGACTA GGGGCAAAAC CGGAACCAAC CACTGAATGG ATGCTGGAA AGACGGTCAG AAATTCACC GTGACCGAG
9601 ATGGCTTGA ATACATATG GGAATCATG AGCCAGTGA GGTCTATGCC CAAGAGTCAG CACCAGGAGA CCTCAGCGA TGGCCACAG AAATAGTACA GCATTACTAC CATCGCCATC
9721 CTGTGTACAC CATCTTAGCC GTCCATCAG CTACCGTGGC GATGATGATT GCGTAAACCG TTGCAGTGTT ATGTGCTGT AAAGCGCGCC GTGAGTGCT GACGCCATAC GCCCTGGCCC
9841 CAAACGCCGT AATCCCACT TCGTGGCAG TCTGTGCTG CTTAGGTGCG GCCAATGCTG AAACGTTAC CGAGACCATG AGTTACTGT GGTGGAACAG TCAGCCGTTT TCTGGGTCC
9961 AGTTGTGCAT ACCTTTGGCC GCTTTCATCG TTCTAATGCG CTGCTGCTCC TGCTGCTGC CTTTTTATG GTTGGCCGC CCTACCTGG CGAAGGTAGA CGCTACGAA CATGCGACCA
10081 CTGTTCCAAA TGTGCCACAG ATACCGTATA AGGCACTGT TGAAGGGCA GGGTATGCC CCGTCAATTT GGAGATCACT GTCATGCTT CCGAGGTTTT GCCTTCCACC AACCAGAGT
10201 ACATTACCTG CAAATTCACC ACTGTGTC CCGCCCAA AATCAAATGC TCGGCTCTT TGAATGTCA GCGGCGCGCT CATGCAACT ATACCTGCAA GGTCTTCGGA GGGGTCTACC
10321 CCTTTATGTG GGGAGGAGCG CAATGTTTT GCGACAGTGA GAACAGCCAG ATGAGTGAGG CGTACGTGA ACTGTACGA GATTGCGCT CTGACCACGC GCAGGCGATT AAGGTGACAA
10441 CTGCGCGAT GAAAGTAGGA CTGCTATAG TGTACGGGA CACTACCAGT TTCTAGATG TGTACGTGA CGGAGTCACA CCAGGAACGT CTAAGAGCTT GAAAGTCATA GCTGGACCAA
10561 TTTCAGCATC GTTACGCCA TTGATCATA AGGTGTTAT CCATCGCGGC CTGGTGATCA ACTATGACTT CCGGAATAT GGAGCGATGA AACCAGGAGC GTTGGAGAC ATTCAGCTA
10681 CCTCCTTAC TAGCAAGGAT CTCATGCCA GCACAGACAT TAGGCTACTC AAGCCTTCC CCAAGAACGT GCATGTCCCG TACACGACAG CCGCATCAGG ATTTGAGATG TGGAAAAACA
10801 ACTCAGGCGC CCACTGACG GAAACCGCAC CTTTCGGTG TAAGATTGCA GTAAATCCCG TCCGAGCGGT GCACTGTTC TACGGGAACA TTCCCATTC TATTGACATC CCGAACGCTG
10921 CTTTATCAG GACATCAGT GCACCACTGG TCTCAACAGT CAAATGTGAA GTGAGTGA GTCACTATTC AGCAGACTT GCGGGGATGG CCACCTGCA GTATGTATCC GACCGGAAG
11041 GTCAATGCC CGTACATTC CATTCGAGCA CAGCAACTCT CCAAGAGTGG ACAGTACATG TCTGGAGAA AGGAGCGGTG ACAGTACACT TTAGCACCCG GAGTCCACAG GCGAACTTTA
11161 TCGTATCGCT GTGTGGGAAG AAGACAACAT GCAATGCAGA ATGTAAACCA CCAGCTGACC ATATCTGAG CACCCCGCAC AAAAATGACC AAGAATTTCA AGCGCCATC TCAAAAACAT
11281 CATGGAGTGG GCTGTTTTCC CTTTCCGGC GCGCTCTG GCTATTAAAT ATAGGACTTA TGATTTTTG TTGACGATG ATGCTGACTA GCACAGGAAG ATGACCGCTA CCGCCCAATG
11401 ATCCGACCA CAAAACCTGA TGTACTCCG AGGAAGTGT GTGCATAATG CATCAGGCTG GTACATTAGA TCCCCGCTTA CCGCGGGCAA TATAGCAACA CTAATAACTC GATGTACTTC
11521 CGAGGAAGCG CAGTGCAATA TGCTGCCAG TTTGCCACA TAACCACTAT ATTAACCAAT TATCTAGCG AGGCCAAAA CTCAATGTAT TTCTGAGGAA GCGTGTGCA TAATGCCAG
11641 CAGCGTCTGC ATAACCTTTA TTATTTCTT TATTAATCAA CAAAATTTG TTTTAAACAT TTC

FIG. 6B